

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: The Trustees of Columbia University in the City of New York City
- (ii) TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS SEQUENCES AND USES THEREOF
- (iii) NUMBER OF SEQUENCES: 45
- (iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Cooper & Dunham LLP
(B) STREET: 1185 Avenue of the Americas
(C) CITY: New York
(D) STATE: New York
(E) COUNTRY: U.S.A.
(F) ZIP: 10036
- (v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.25
- (vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: White, John P.
(B) REGISTRATION NUMBER: 28,678
(C) REFERENCE/DOCKET NUMBER: 45185-C-PCT/JPW/MS
- (ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (212) 278-0400
(B) TELEFAX: (212) 391-0525

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20710 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: N
- (iv) ANTI-SENSE: N
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
- | | |
|---|-----|
| TCGAGTCGGA GAGTTGGCAC AGGCCTTGAG CTCGCTGTGA CGTTCTCACG GTGTTGGTTG | 60 |
| GGATCAGCTG GTGACTCAGA CAAGTCTTGA GCTCTACAAC GTAACATACG GGCTGATGCC | 120 |
| CACCCGATAC CAGAATTACG CAGTCGGCAA TTCTGTGCCC TAGAGTCACC TCAAAGAATA | 180 |
| ATCTGTGGTG TCCAAGGGGA GGGTTCTGGG GCCGGCTACT TAGAAACCGC CATAGATCGG | 240 |

GCAGGGGTGGA	GTACTTGAGG	AGCCGGCCGT	AGGTGGCCAG	GTGGGCCCGG	TTACCTGCTC	300
TTTTGCGTGC	TGCTGGAAGC	CTGCTCAGGG	ATTTCTTAAC	CTCGGCCTCG	GTTGGACGTA	360
CCATGGCAGA	AGGCGGTTTT	GGAGCGGACT	CGGTGGGGCG	CGGCGGAGAA	AAGGCCTCTG	420
TGACTAGGGG	AGGCAGGTGG	GACTTGGGGA	GCTCGGACGA	CGAATCAAGC	ACCTCCACAA	480
CCAGCACGGA	TATGGACGAC	CTCCCTGAGG	AGAGGAAACC	ACTAACGGGA	AAGTCTGTAA	540
AAACCTCGTA	CATATACGAC	GTGCCCACCG	TCCCGACCAG	CAAGCCGTGG	CATTTAATGC	600
ACGACAACCT	CCTCTACGCA	AGGCCTAGGT	TTCCGCCCAG	ACCTCTCATA	CGGCACCCCTT	660
CCGAAAAAGG	CAGCATTTTT	GCCAGTCGGT	TGTCAGCGAC	TGACGACGAC	TCGGGAGACT	720
ACGCGCCAAT	GGATCGCTTC	GCCTTCCAGA	GCCCCAGGGT	GTGTGGTGGC	CCTCCCTTC	780
CGCCTCCAAA	TCACCCACCT	CCGGCAACTA	GGCCGGCAGA	CGCGTCAATG	GGGGACGTGG	840
GCTGGGCGGA	TCTGCAGGGA	CTCAAGAGGA	CCCCAAAGGG	ATTTTTAAAA	ACATCTACCA	900
AGGGGGGCGAG	TCTCAAAGCC	CGTGGACGCG	ATGTAGGTGA	CCGTCTCAGG	GACGGCGGCT	960
TTGCCTTTAG	TCCTAGGGGC	GTGAAATCTG	CCATAGGGCA	AAACATTAAA	TCATGGTTGG	1020
GGATCGGAGA	ATCATCGGCG	ACTGCTGTCC	CCGTCAACAC	GCAGCTTATG	GTACCGGTGC	1080
ACCTCATTAG	AACGCCTGTG	ACCGTGGACT	ACAGGAATGT	TTATTTGCTT	TACTTAGAGG	1140
GGGTAATGGG	TGTGGGCAAA	TCAACGCTGG	TCAACGCCGT	GTGCGGGATC	TTGCCCCAGG	1200
AGAGAGTGAC	AAGTTTTCCC	GAGCCCATGG	TGTACTGGAC	GAGGGCATT	ACAGATTGTT	1260
ACAAGGAAAT	TTCCACCTG	ATGAAGTCTG	GTAAGGCGGG	AGACCCGCTG	ACGTCTGCCA	1320
AAATATACTC	ATGCCAAAAC	AAGTTTTTCG	TCCCCTTCCG	GACGAACGCC	ACCGCTATCC	1380
TGCGAATGAT	GCAGCCCTGG	AACGTTGGGG	GTGGGTCTGG	GAGGGGCACT	CACTGGTGCG	1440
TCTTTGATAG	GCATCTCCTC	TCCCAGCAG	TGGTGTTCCC	TCTCATGCAC	CTGAAGCACG	1500
GCCGCCTATC	TTTTGATCAC	TTCTTCAAT	TACTTTCCAT	CTTTAGAGCC	ACAGAAGGCG	1560
ACGTGGTCGC	CATTCTCACC	CTCTCCAGCG	CCGAGTCGTT	GCGGCGGGTC	AGGGCGAGGG	1620
GAAGAAAGAA	CGACGGGACG	GTGGAGCAAA	ACTACATCAG	AGAATTGGCG	TGGGCTTATC	1680
ACGCCGTGTA	CTGTTTCATG	ATCATGTTGC	AGTACATCAC	TGTGGAGCAG	ATGGTACAAC	1740
TATGCGTACA	AACCACAAAT	ATTCCGGAAA	TCTGCTTCCG	CAGCGTGCGC	CTGGCACACA	1800
AGGAGGAAAC	TTTGAAAAAC	CTTCACGAGC	AGAGCATGCT	ACCTATGATC	ACCGGTGTAC	1860
TGGATCCCGT	GAGACATCAT	CCCGTCGTGA	TCGAGCTTTG	CTTTTGTTTC	TTCACAGAGC	1920
TGAGAAAATT	ACAATTTATC	GTAGCCGACG	CGGATAAGTT	CCACGACGAC	GATGCGGCC	1980
TGTGGACCGA	AATCTACAGG	CAGATCCTGT	CCAATCCGGC	TATTAAACCC	AGGGCCATCA	2040
ACTGGCCAGC	ATTAGAGAGC	CAGTCTAAAG	CAGTTAATCA	CCTAGAGGAG	ACATGCAGGG	2100
TCTAGCCTTC	TTGGCGGCCC	TTGCATGCTG	GCGATGCATA	TCGTTGACAT	GTGGAGCCAC	2160
TGGCGCGTTG	CCGACAACGG	CGACGACAAT	AACCCGCTCC	GCCACGCAGC	TCATCAATGG	2220
GAGAACCAAC	CTCTCCATAG	AACTGGAATT	CAACGGCACT	AGTTTTTTTC	TAAATTGGCA	2280

AAATCTGTTG AATGTGATCA CGGAGCCGGC CCTGACAGAG TTGTGGACCT CCGCCGAAGT 2340
CGCCGAGGAC CTCAGGGTAA CTCTGAAAAA GAGGCAAAGT CTTTTTTTCC CCAACAAGAC 2400
AGTTGTGATC TCTGGAGACG GCCATCGCTA TACGTGCGAG GTGCCGACGT CGTCGCAAAC 2460
TTATAACATC ACCAAGGGCT TTAACATAG CGCTCTGCCC GGGCACCTTG GCGGATTG 2520
GATCAACGCG CGTCTGGTAC TGGGTGATAT CTTGCGCATCA AAATGGTCGC TATTCGCGAG 2580
GGACACCCCA GAGTATCGGG TGTTTTACCC AATGAATGTC ATGGCCGTCA AGTTTTCCAT 2640
ATCCATTGGC AACAACGAGT CCGGCGTAGC GCTCTATGGA GTGGTGTGCG AAGATTTCGT 2700
GGTCGTCACG CTCCACAACA GGTCACAAAGA GGCTAACGAG ACGGCGTCCC ATCTTCTGTT 2760
CGGTCTCCCG GATTCACCTG CATCTCTGAA GGGCCATGCC ACCTATGATG AACTCACGTT 2820
CGCCCGAAAC GCAAAATATG CGCTAGTGGC GATCCTGCCT AAAGATTCTT ACCAGACACT 2880
CCTTACAGAG AATTACACTC GCATATTTCT GAACATGACG GAGTCGACGC CCCTCGAGTT 2940
CACGCGGACG ATCCAGACCA GGATCGTATC AATCGAGGCC AGGCGCGCCT GCGCAGCTCA 3000
AGAGGCGGCG CCGGACATAT TCTTGGTGTG GTTTCAGATG TTGGTGGCAC ACTTTCTTGT 3060
TGCGCGGGGC ATTGCCGAGC ACCGATTGTG GAGGTGGAC TCGTGTGTC GGCAGTATGC 3120
GGAAGTGTAT TTTCTCCGCC GCATCTCGCG TCTGTGCATG CCCACGTTCA CCACTGTGCG 3180
GTATAACCAC ACCACCCTTG GCGCTGTGGC CGCCACACAA ATAGCTCGCG TGTCCGCCAC 3240
GAAGTTGGCC AGTTTGCCCC GCTCTTCCCA GGAAACAGTG CTGGCCATGG TCCAGCTTGG 3300
CGCCCGTGAT GCGCGCGTCC CTTCTCCAT TCTGGAGGCG ATTGCTATGG TCGTCGAACA 3360
TATGTATACC GCCTACACTT ATGTGTACAC ACTCGGCGAT ACTGAAAGAA AATTAATGTT 3420
GGACATACAC ACGGTCCTCA CCGACAGCTG CCCGCCCAA GACTCCGGAG TATCAGAAAA 3480
GCTACTGAGA ACATATTTGA TGTTACATC AATGTGTACC AACATAGAGC TGGGCGAAAT 3540
GATCGCCCGC TTTTCCAAAC CGGACAGCCT TAACATCTAT AGGGCATTCT CCCCCTGCTT 3600
TCTAGGACTA AGGTACGATT TGCATCCAGC CAAGTTGCGC GCCGAGGCGC CGCAGTCGTC 3660
CGCTCTGACG CGGACTGCCG TTGCCAGAGG AACATCGGGA TTCGCAGAAAT TGCTCCACGC 3720
GCTGCACCTC GATAGCTTAA ATTTAATTCC GCGGATTAAC TGTTCAAAGA TTACAGCCGA 3780
CAAGATAATA GCTACGGTAC CTTGCGCTCA CGTCACGTAT ATCATCAGTT CCGAAGCACT 3840
CTCGAACGCT GTTGTCTACG AGGTGTGCGA GATCTTCCTC AAGAGTGCCA TGTATTATATC 3900
TGCTATCAAA CCCGATTGCT CCGGCTTTAA CTTTCTCAG ATTGATAGGC ACATTCCCAT 3960
AGTCTACAAC ATCAGCACAC CAAGAAGAGG TTGCCCCCTT TGTGACTCTG TAATCATGAG 4020
CTACGATGAG AGCGATGGCC TGCAGTCTCT CATGTATGTC ACTAATGAAA GGGTGCAGAC 4080
CAACCTCTTT TTAGATAAGT CACCTTTCTT TGATAATAAC AACCTACACA TTCATTATT 4140
GTGGCTGAGG GACAACGGGA CCGTAGTGGA GATAAGGGGC ATGTATAGAA GACGCGCAGC 4200
CAGTGCTTTG TTTCTAATTC TCTCTTTTAT TGGGTTCTCG GGGGTTATCT ACTTTCTTTA 4260
CAGACTGTTT TCCATCCTTT ATTAGACGGT CAATAAAGCG TAGATTTTTA AAAGGTTTCC 4320

TGTCATTCT	TTTGTATGG	GCATATACTT	GGCAAGAAAT	CCGAGCACCT	CAGAAAGTGG	4380
ATTGCCGTCA	CATATCAGTT	CGACCACCCC	TGCACCTAGC	CATGCGGCGC	TTTGACGGTC	4440
TTTGGGGCTA	CACATCATAA	AGTACTTTTC	CATGGCTTCT	ATAAGCACCT	TGGAACAATC	4500
TGGGGGTTGG	CGAATGGGTT	CCCTAAACGG	GAAATCCTCT	ATGGTATTCA	GGCAGAAGAC	4560
CGCGTCTCTC	ACCCGACGTT	TGAGTCTTTC	TAGCAGAGCG	CCGAAGAACT	CCCGCTCGTG	4620
TGTTTTTCGCA	GGGGCAAGTT	CTGCGCCGTA	CAGCGATGAG	AAACACGACA	CGATGTTTTTC	4680
CAGCCCCATG	CTGCGCAGCA	ACACGTGCTT	CAGGAACAGG	TGTTGTAGCC	GGTTCAGTTT	4740
TAGCTTGGGT	AGAAAAGTTA	TCGAGTTGTT	AGCACGCTCC	ATGATGGTAA	CGGTGTTGAA	4800
GTCACAGACC	GGGCTTTCTC	CGAGTCTCGG	CCGCTGAGT	CCAATCATGT	AGAACATAGA	4860
CGCGGCCTCG	TTGTCTGTGT	TAAGTGACAC	GATATCCCGT	TCGCAAACCT	GTGCGATGTT	4920
GTGTTTCAGT	ATAGATCTGG	TCTGACCGGC	ACGGGGTGTT	ATGGGGTGAC	GCGGTAAAGG	4980
CGACTCTGGG	TCAAACACCT	TTATGCGGTT	GGCGGCCTCG	TCGATGACGA	CACGCTTGTT	5040
CGCGGCGTGT	ATGGGGACGC	GACGGCATCC	CGCTGGCAGA	TCTATAATCT	TAAAGTTGGT	5100
ATAAGACTGG	TCGCTCGTTA	TGGCCAGCCG	GCACTCCGGT	AGTATCTGCG	TGTCTCGAA	5160
TTCGTGGCCG	CGTACGACTG	GCTTGGAGTG	CAGGTAAACG	CCAAGAGATG	CGGTCTCTTC	5220
GCCTACGCAC	AAGTGGCTTC	TTAACGCGTA	GGGGTCCGGT	GAGAGCATGA	TCCGTAGCAA	5280
CGATAGTTCC	GGGTGCCTAG	CCGCGTAGAG	TGGCAGCGTA	GACGAGTCCG	GAGTCCCCAA	5340
CTTTTCGAAC	AACAGTGGCA	TCGGGACTTC	AGGATTAGAG	ACTCCCACCA	TGGCCGCCAC	5400
CGCCGGAGAG	GTCAAGACGT	GAAACACGCG	CTCGCCTGTC	GACAGGCGCG	CCGCGCCCTC	5460
TACTAGACTA	GCCTTCACGT	CCGGAACCTG	TAACATAGCT	TAGACCAGCG	GACGGACGCA	5520
ACGTACGCGG	GGATCGGCTG	GCGGTGTCTG	CTCGTTGGAC	GCGGCCGTTT	GGTGGCGCCA	5580
GTGCAGGCCT	AGTTTGC GAA	TGGCGTGACG	GACAATTTGT	GGCTTTAGAG	CGGCGAACCG	5640
ATGACCCGTG	GTGGCGACGA	ACGAAATGAA	GTTTGCATTG	CGGCCAACT	CGTCTAGCCT	5700
GGTCTTCTTG	TTTCGGGCAT	AGATTTTCGG	GATTAGGTTA	CACTTTTAT	ATCCCAGTAC	5760
TGCGCACTCG	TGTTTGCTTT	TAGTGTGACT	GATTATCTTC	TTTGAGAGT	CAAACAGGCC	5820
CCGGGCGGCG	GCTCGCCTAA	TGCAAGCCAC	GTCAAGCCTG	AGAAACGAAC	AGCATTCCAC	5880
CAGACACTCC	AGGAACCTTT	TGTGTAGCGT	CTGTATTTGG	GAACGGTTTC	TGTGCTCAAG	5940
TAGGGAGAAT	ATTCTATTTT	TGTTTCCGTC	GATGCGCGCG	TGCTGGTCCG	TGAGAATGGG	6000
CGCCAGCTCG	TGGCGAATCT	GTTCCACAAG	AGGCTGCCCC	TACACTTTAG	AAATCGTGGC	6060
TGTCGCGGCC	TTAAACCAGG	ACACGTTTAG	CCCATCCTTG	CTGGAGACCA	CAGATGGAAA	6120
GTTTGTGGTC	CAAAATACGT	TTTTTCGCC	CATTCTCACC	ATGTACTGGT	TTTCCAGTCC	6180
GTGCAGGTCC	AACGTGGAGT	TCCAATTTGC	TATCGATACA	GGAAATATGT	GCCTGATTGG	6240
CAGAAAGCAT	TTCAGCGTAC	CCATTGCGAA	GAGAAAGTGC	AGCATGTCCC	CACTGATGTT	6300
GATGTTTATT	GCGGTGCCTT	GACACATGTT	GTCGGAAAAA	AACACGCTTA	TGGTAAAAGA	6360

AGGTTCTTTT ACGGAGTACT TTCGTATAAC AAAATTGTTG GTCAATCTGG GGATGTTTAA 6420
AATAGTCTTT TGCAGGGTGT TAGGAACGTG GCAGCTTATC TTAGTGTTAA TCACCATGTT 6480
GGTGTGAAT ATGGTGATCT TGAAGTTTTT CAAACTGACG TGTTTTGTGG GTTCCAGCAT 6540
GTCTGACACT GTAGAGCTGC CCAGAGTCCG CGCGTCCGTG GCCGCGTATC GTTGAAGCA 6600
CGCCTGCAAA TTTCCTTTCA TGGCTGCTCG CCGGTCTTTC GGCGCGTACC GGATTCTTGA 6660
AAGCGTCGCC GCCAGGAGAC GCGGTGTCTC GTGGGTGCCT AAAAAGTTTG CGCAGGGGTG 6720
CAGTCCGCTG CACGAGTGGC CGATGCAGTC TGCCACTGCC ATACACATGA CGAGTCTGTA 6780
GATGGCCGGT GTGCCCGGAT AACTAGATA GTAGGTACAA TCTGGGGTAC TGACGACCAC 6840
CCTGTATGGC TTTGGTCCGG GGTCTTTCG TTGGATTTTT ACGTGCAGAC GGGACACGAG 6900
CTGGTTTAGA GCCAGCTGAA AGCCCACCAG ATCCCGTCCG TTAACCTTGA CGTCTGGTG 6960
CTTACTCTGT TTCGACAGGT TCTTCAGCAC GGTGGGCAGT CGCTCTACGT TGTGAGCGAT 7020
GGCAGGCGC AGCGAGACCA GCTCTCCGTG CCACCCCCAC GTGGCCATGA AGCTGCTGAT 7080
GTTAACTTT AAAAATGTA GCTGTGCGTC TGGGGATGCG GGTGGCATT TGAACACGA 7140
GAGATGCTT AGGCTCTCCA GGAGTGCAAA ATAATTTTGA TAGATTGTGG GTTGTAGACT 7200
ATGGGGCAAC ACCGCCAGAA ACGCATGAAA ACCTGTTCG AACTCCCAGA ACTCCAGGTA 7260
CCTGCACACT ATCCTGAACA TGGCTTTGTA ACATATGGTG CACGTTAGTA GCGCGGGAAG 7320
ATACAGCGAG CGTAGCTCCC TGAATTCGCA GGGTTTATCA CAATCATCGG TAAGTTCCCA 7380
TGATCCCACC GCAGGTAGGT AGTTGTCGGT GTCTATCTGT CCGCGCGTAA AACTCCACC 7440
ACCGTCAATT ATTAAACCTT CGCCGCTGTA CCGTCGACCC ACTTTTCCCA AAAGAGTCCC 7500
TTCTTGATGT ATAAAAGGGT GGAGGCGTTC CCCCAGGAGT AGTCTGCGTA TCGCTCTGCA 7560
GGCGAAAAAG GTGGGCTCGG GCTGCATCAT CTTATCAAGA CCTTCTAAGG TCAGCTCTGC 7620
CTGCAGGTGC GAGTTGGTGG CCAGACAGCA GAATATTTCC AGCTGTGATT CCCAAGTCGC 7680
TTGATAACAC GTGGTCTGCG GACTCGTCGT CAGGGAGGCG CTCGTTGGCA GTAGTAGGGG 7740
GCCCTCGAGC GCTGCCATGG AGGCGACCTT GGAGCAACGA CCTTTCCCGT ACCTCGCCAC 7800
GGAGGCCAAC CTCCTAACGC AGATTAAAGGA GTCGGCTGCC GACGGACTCT TCAAGAGCTT 7860
TCAGCTATTG CTCGGCAAGG ACGCCAGAGA AGGCAGTGTG CGTTTCGAAG CGCTACTGGG 7920
CGTATATACC AATGTGGTGG AGTTTGTTAA GTTCTGGAG ACCGCCCTCG CCGCCGCTTG 7980
CGTCAATACC GAGTTCAAGG ACCTGCGGAG AATGATAGAT GGAAAAATAC AGTTTAAAT 8040
TTCAATGCCC ACTATTGCCC ACGGAGACGG GAGGAGGCCC AACAAGCAGA GACAGTATAT 8100
CGTCATGAAG GCTTGCAATA AGCACCACAT CCGTGCGGAG ATTGAGCTTG GGGCCGAGA 8160
CATCGAGCTT CTCTTCGCCG AGAAAGAGAC GCCCTTGGAC TTCACAGAGT ACGCGGGTGC 8220
CATCAAGACG ATTACGTCGG CTTTGCAATT TGGTATGGAC GCCCTAGAAC GGGGCTAGT 8280
GGACACGGTT CTCGCAGTTA AACTTCGGCA CGTCCACCC GTCTTTATTT TAAAGACGCT 8340
GGGCGATCCC GTCTACTCTG AGAGGGGCCT CAAAAGGCC GTCAAGTCTG ACATGGTATC 8400

CATGTTCAAG GCACACCTCA TAGAACATTC ATTTTTTCTA GATAAGGCCG AGCTCATGAC 8460
AAGGGGGAAG CAGTATGTCC TAACCATGCT CTCCGACATG CTGGCCGCGG TGTGCGAGGA 8520
TACCGTCTTT AAGGGTGTCA GCACGTACAC CACGGCCTCT GGGCAGCAGG TGGCCGGCGT 8580
CCTGGAGACG ACGGACAGCG TCATGAGACG GCTGATGAAC CTGCTGGGGC AAGTGGAAGG 8640
TGCCATGTCC GGGCCCGCGG CCTACGCCAG CTACGTTGTC AGGGGTGCCA ACCTCGTCAC 8700
CGCCGTTAGC TACGGAAGGG CGATGAGAAA CTTTGAACAG TTTATGGCAC GCATAGTGGA 8760
CCATCCCAAC GCTCTGCCGT CTGTGGAAGG TGACAAGGCC GCTCTGGCGG ACGGACACGA 8820
CGAGATTCAG AGAACCCGCA TCGCCGCCTC TCTCGTCAAG ATAGGGGATA AGTTTGTGGC 8880
CATTGAAAGT TTGCAGCGCA TGTACAACGA GACTCAGTTT CCCTGCCCCAC TGAACCGGCG 8940
CATCCAGTAC ACCTATTTCT TCCCTGTTGG CCTTCACCTT CCCGTGCCCC GCTACTCGAC 9000
ATCCGTCTCA GTCAGGGGCG TAGAATCCCC GGCCATCCAG TCGACCGAGA CGTGGGTGGT 9060
TAATAAAAAC AACGTGCCTC TTTGCTTCGG TTACCAAAC GCCCTCAAAA GCATATGCCA 9120
CCCTCGAATG CACAACCCCA CCCAGTCAGC CTAGGCACTA AACCAAGCTT TTCCCGATCC 9180
CGACGGGGGA CATGGGTACG GTCTCAGGTA TGAGCAGACG CCAAACATGA ACCTATTCAG 9240
AACGTTCCAC CAGTATTACA TGGGGAAAAA CGTGGCATTG GTTCCCGATG TGGCCCCAAA 9300
AGCGCTCGTA ACCACGGAGG ATCTACTGCA CCCAAGCTCT CACCGTCTCC TCAGATTGGA 9360
GGTCCACCCC TTCTTTGATT TTTTGTGCA CCCCTGTCCT GGAGCGAGAG GATCGTACCG 9420
CGCCACCCAC AGAACAATGG TTGGAAATAT ACCACAACCG CTCGCTCCAA GGGAGTTTCA 9480
GGAAAGTAGA GGGGCGCAGT TCGACGCTGT GACGAATATG ACACACGTCA TAGACCAGCT 9540
AACTATTGAC GTCATACAGG AGACGGCATT TGACCCCGCG TATCCCCTGT TCTGCTATGT 9600
AATCGAAGCA ATGATTACAG GACAGGAAGA AAAATTCGTG ATGAACATGC CCCTCATGTC 9660
CCTGGTCATT CAAACCTACT GGGTCAACTC GGGAAACTG GCGTTTGTGA ACAGTTATCA 9720
CATGGTTAGA TTCATCTGTA CGCATATTGG GAATGGAAGC ATCCCTAAGG AGGCGCACGG 9780
CCACTACCGG AAAATCTTAG GCGAGCTCAT CGCCCTTGAG CAGGCGCTTC TCAAGCTCGC 9840
GGGACACGAG ACGGTGGGTC GGACGCCGAT CACACATCTG GTTTCGGCTC TCCTCGACCC 9900
GCATCTGCTG CCTCCCTTTG CCTACCACGA TGTCTTTACG GATCTTATGC AGAAGTCATC 9960
CAGACAACCC ATAATCAAGA TCGGGGATCA AAATACGAC AACCCTCAA ATAGGGCGAC 10020
ATTCATCAAC CTCAGGGGTC GCATGGAGGA CCTAGTCAAT AACCTTGTTA ACATTTACCA 10080
GACAAGGGTC AATGAGGACC ATGACGAGAG ACACGTCCTG GACGTGGCGC CCTGGACGA 10140
GAATGACTAC AACC CGGTCC TCGAGAAGCT ATTCTACTAT GTTTTAATGC CGGTGTGCAG 10200
TAACGGCCAC ATGTGCGGTA TGGGGGTCGA CTATCAAAC GTGGCCCTGA CGCTGACTTA 10260
CAACGGCCCC GTCTTTGCGG ACGTCGTGAA CGCACAGGAT GATATTCTAC TGCACCTGGA 10320
GAACGGAACC TTGAAGGACA TTCTGCAGGC AGGCGACATA CGCCCGACGG TGGACATGAT 10380
CAGGGTGCTG TGCACCTCGT TTCTGACGTG CCCTTTCGTC ACCCAGGCCG CTCGCGTGAT 10440

00507479-060000

CACAAAGCGG	GACCCGGGCC	AGAGTTTTC	CACGCACGAA	TACGGGAAGG	ATGTGGCGCA	10500
GACCGTGCTT	GTAAATGGGT	TTGGTGCGTT	CGCGGTGGCG	GACCGCTCTC	GCGAGGCGGC	10560
GGAGACTATG	TTTTATCCGG	TACCCTTTAA	CAAGCTCTAC	GCTGACCCGT	TGGTGGCTGC	10620
CACACTGCAT	CCGCTCCTGC	CAAACTATGT	CACCAGGCTC	CCCAACCAGA	GAAACGCGGT	10680
GGTCTTTAAC	GTGCCATCCA	ATCTCATGGC	AGAATATGAG	GAATGGCACA	AGTCGCCCCG	10740
CGCGGCGTAT	GCCGCGTCTT	GTCAGGCCAC	CCCGGGCGCC	ATTAGCGCCA	TGGTGAGCAT	10800
GCACCAAAAA	CTATCTGCCC	CCAGTTTCAT	TTGCCAGGCA	AAACACCGCA	TGCACCCTGG	10860
TTTTGCCATG	ACAGTCGTCA	GGACGGACGA	GGTTCTAGCA	GAGCACATCC	TATACTGCTC	10920
CAGGGCGTCG	ACATCCATGT	TTGTGGGGCT	GCCTTCGGTG	GTACGGCGCG	AGGTACGTTT	10980
GGACGCGGTG	ACTTTTGAAA	TTACCCACGA	GATCGCTTCC	CTGCACACCG	CACTTGGCTA	11040
CTCATCAGTC	ATCGCCCCGG	CCCACGTGGC	CGCCATAACT	ACAGACATGG	GAGTACATTG	11100
TCAGGACCTC	TTTATGATTT	TCCCAGGGGA	CCCGTATCAG	GACCGCCAGC	TGCATGACTA	11160
TATCAAAATG	AAAGCGGGCG	TGCAAACCGG	CTCACC GGGA	AACAGAATGG	ATCACGTGGG	11220
ATACACTGCT	GGGGTTCCTC	GCTGCGAGAA	CCTGCCCCGT	TTGAGTCATG	GTCAGCTGGC	11280
AACCTGCGAG	ATAATTCCCA	CGCCGGTCAC	ATCTGACGTT	GCCTATTTCC	AGACCCCCAG	11340
CAACCCCCGG	GGGCGTGCGG	CGTCGGTCGT	GTCGTGTGAT	GCTTACAGTA	ACGAAAGCGC	11400
AGAGCGTTTG	TTCTACGACC	ATTCAATACC	AGACCCCGCG	TACGAATGCC	GGTCCACCAA	11460
CAACCCGTGG	GCTTCGCAGC	GTGGCTCCCT	CGGCGACGTG	CTATACAATA	TCACCTTTTC	11520
CCAGACTGCG	CTGCCGGGCA	TGTACAGTCC	TTGTGCGCAG	TTCTTCCACA	AGGAAGACAT	11580
TATGCGGTAC	AATAGGGGGT	TGTACACTTT	GGTTAATGAG	TATCTTGCCA	GGCTTGCTGG	11640
GGCCCCCGCC	ACCAGCACTA	CAGACCTCCA	GTACGTCGTG	GTCAACGGTA	CAGACGTGTT	11700
TTTGGACCAG	CCTTGCCATA	TGCTGCAGGA	GGCCTATCCC	ACGCTGCGCG	CCAGCCACAG	11760
AGTTATGCTT	GCCGAGTACA	TGTCAAACAA	GCAGACACAC	GCCCCAGTAC	ACATGGGGCA	11820
GTATCTCATT	GAAGAGGTGG	CGCCGATGAA	GAGACTATTA	AAGCTCGGAA	ACAAGGTGGT	11880
GTATTAGCTA	ACCCTTCTAG	CGTTGGCTAG	TCATGGCACT	CGACAAGAGT	ATAGTGGTTA	11940
ACTTCACCTC	CAGACTCTTC	GCTGATGAAC	TGGCCGCCCT	TCAGTCAAAA	ATAGGGAGCG	12000
TACTGCCGCT	CGGAGATTGC	CACCGTTTAC	AAAATATACA	GGCATTGGGC	CTGGGGTGCG	12060
TATGCTCACG	TGAGACATCT	CCGGACTACA	TCCAAATTAT	GCAGTATCTA	TCCAAGTGCA	12120
CACTCGCTGT	CCTGGAGGAG	GTTGCGCCGG	ACAGCCTGCG	CCTAACGCGG	ATGGATCCCT	12180
CTGACAACCT	TCAGATAAAA	AACGTATATG	CCCCCTTTTT	TCAGTGGGAC	AGCAACACCC	12240
AGCTAGCAGT	GCTACCCCCA	TTTTTTAGCC	GAAAGGATTC	CACCATTTGT	CTCGAATCCA	12300
ACGGATTTGA	CCCCGTGTTT	CCCATGGTCG	TGCCGCAGCA	ACTGGGGCAC	GCTATTCTGC	12360
AGCAGCTGTT	GGTGTACCAC	ATCTACTCCA	AAATATCGGC	CGGGGCCCCG	GATGATGTAA	12420
ATATGGCGGA	ACTTGATCTA	TATACCACCA	ATGTGTCATT	TATGGGGCGC	ACATATCGTC	12480

TGGACGTAGA	CAACACGGAT	CCACGTA CTG	CCCTGCGAGT	GCTTGACGAT	CTGTCCATGT	12540
ACCTTTGTAT	CCTATCAGCC	TTGGTTCCCA	GGGGGTGTCT	CCGTCTGCTC	ACGGCGCTCG	12600
TGCGGCACGA	CAGGCATCCT	CTGACAGAGG	TGTTTGAGGG	GGTGGTGCCA	GATGAGGTGA	12660
CCAGGATAGA	TCTCGACCAG	TTGAGCGTCC	CAGATGACAT	CACCAGGATG	CGCGTCATGT	12720
TCTCCTATCT	TCAGAGTCTC	AGTTCTATAT	TTAATCTTGG	CCCCAGACTG	CACGTGTATG	12780
CCTACTCGGC	AGAGACTTTG	GCGGCCTCCT	GTTGGTATTC	CCCACGCTAA	CGATTTGAAG	12840
CGGGGGGGGT	ATGGCGTCAT	CTGATATTCT	GTCGGTTGCA	AGGACGGATG	ACGGCTCCGT	12900
CTGTGAAGTC	TCCCTGCGTG	GAGGTAGGAA	AAAAACTACC	GTCTACCTGC	CGGACACTGA	12960
ACCCTGGGTG	GTAGAGACCG	ACGCCATCAA	AGACGCCTTC	CTCAGCGACG	GGATCGTGGA	13020
TATGGCTCGA	AAGCTTCATC	GTGGTGCCCT	GCCCTCAAAT	TCTCACAACG	GCTTGAGGAT	13080
GGTGCTTTTT	TGTTATTGTT	ACTTGCAAAA	TTGTGTGTAC	CTAGCCCTGT	TTCTGTGCCC	13140
CCTTAATCCT	TACTTGGTAA	CTCCCTCAAG	CATTGAGTTT	GCCGAGCCCG	TTGTGGCACC	13200
TGAGGTGCTC	TTCCACACACC	CGGCTGAGAT	GTCTCGCGGT	TGCGATGACG	CGATTTTCTG	13260
TAAACTGCCC	TATACCGTGC	CTATAATCAA	CACCAAGTTT	GGACGCATTT	ACCCGAACTC	13320
TACACGCGAG	CCGGACGGCA	GGCCTACGGA	TTACTCCATG	GCCCTTAGAA	GGGCTTTTGC	13380
AGTTATGGTT	AACACGTCAT	GTGCAGGAGT	GACATTGTGC	CGCGGAGAAA	CTCAGACCGC	13440
ATCCCCTAAC	CACACTGAGT	GGGAAAATCT	GCTGGCTATG	TTTTCTGTGA	TTATCTATGC	13500
CTTAGATCAC	AACTGTCACC	CGGAAGCACT	GTCTATCGCG	AGCGGCATCT	TTGACGAGCG	13560
TGACTATGGA	TTATTCACTC	CTCAGCCCCG	GAGCGTGCCC	TCGCTACCC	CTTGCGACGT	13620
GTCGTGGGAA	GATATCTACA	ACGGGACTTA	CCTAGCTCGG	CCTGGAAACT	GTGACCCCTG	13680
GCCCAATCTA	TCCACCCCTC	CCTTGATTCT	AAATTTTAAA	TAAAGGTGTG	TCACTGGTTA	13740
CACCACGATT	AAAAACCACT	CACTGAGATG	TCTTTTAAAC	CGCTAAGGGA	TTATACCGGG	13800
ATTTAAAACC	GCCCACTGAT	TTTTTTACGC	TAAGAGTTGG	GTGCTTGGGG	GGTTTTGCAT	13860
TGCTCTGTTG	TAAACTATAT	ATAAGTTAAA	CCAAAATTCT	CAGGGAGACA	AGGTGACGGT	13920
GGTGAGAACT	CAGTTGAGAG	TCAGAGAATA	CAGTGCTAAT	CAGGGTAGAT	GAGCATGACT	13980
TTCCCCGTCT	CCAGTCACCG	GAGGAATGGT	GGACGGCTCC	GTCCTGGTGC	GAATGGCCAC	14040
CAAGCCTCCC	GTGATTGGTC	TTATAACAGT	GCTCTTCCTC	CTAGTCATAG	GCGCCTGCGT	14100
CTACTGCTGC	ATTGCGGTGT	TCCTGGCGGC	TCGACTGTGG	CGCGCCACCC	CACTAGGCAG	14160
GGCCACCGTG	GCGTATCAGG	TCCTTCGCAC	CCTGGGACCG	CAGGCCGGGT	CACATGCAAC	14220
GCCGACGGTG	GGCATAGCTA	CCCAGGAGCC	CTACCGTACA	ATATACATGC	CAGATTAGAA	14280
CGGGGTGTGT	GCTATAATGG	ATGGCTATGG	GGGGGGGCTG	TAGATAATTG	AGCGCTGTGC	14340
TTTTATTGTG	GGGATATGGG	CTTGATACATG	TGTCTATCAT	CGGTAGCCAT	AAAATGGGCC	14400
ATGACAACCTG	CCACAAGTAA	GTCGTCCGAC	ATGTGCTTTT	GCTTGGCGCT	GTATGACTGC	14460
CCTCCATCCC	TAAGCGGGAC	GCACTTGATC	GCGCGGACCT	GTTCTACCAG	GTAGGTCACC	14520

GGGTCAAATG ATATTTTGAT GGTGTTGGAC ACCACCGTCT GGCTGGCGCT CAGGGTGCCG 14580
GAGTTCAGAG CGTAGATGAA TGTCTCAAAC GCGGAGGATT TCTCGCCTCC CAACATGTAA 14640
ATTGGCCACT GCAGGGCGCT GCTCTTGTC A GTATAGTGTA GAAAATGTAT GGGGAGCGGG 14700
CATATTTTCGT TAAGGACGGT TGCAATGGCC ACCCCAGAAT CTTGGCTGCT GTTGCCCTTCG 14760
ACCGCCGCGT TCACGCGCTC AATTGTGGTG TGGAGCACAG CGATCGCCTT AATCATCGTG 14820
CATGCGCAGG ACGCTATCTC GTAAGCAGCT GCGCCAGTGA GGTCGCGCAG GAAGAAATGC 14880
TCCATGCCCA ATATGAGGCT TCTGGTGGGA GTCTGAGTAC TCGTGACAAC GGCGCCCACG 14940
CCAGTACCGG ACGCCTCCGT GTTGTTTCGTA TACGCGGGGT CGATGTAAAC AAACAGCTGT 15000
TTTCCAAGGC ACTTCTGAAC CTCCTGGGCG GTGGTGTCTA CCCGACACAT GTCAAACCTGT 15060
GTCAGCGCTG CGTCACCCAC CACGCGGTAA AGCGTAGCAT TTGACGACGC TGCTCCCTCG 15120
CCCATTAGTT CGGTGTCGAA TGCCCCCTCC ATAAAGAGGT TGGTGGTGGT TTTGATGGAT 15180
TCGTCGATGG TGATGTACGT CGGAATGTGC AGTCTGTAC AAGGACAGGA CACTAGTGCG 15240
TCTTGACAGT GGAAATCTTC TCGGTGGTCC GCACACACGT AACTGACCAC ATTCAGCATC 15300
TTTTCTGGG CGTTCCTGAG GTTAAGCAGG AACTCTGTGG AGCGGTCTGA CGAGTTCACG 15360
GATGATATAA ATATAAGCTT GGCCTCTTTC TGAAGCATGA AACCAGAAT AGCCGGCAGT 15420
GCATCCTTTT TAATAAAATT CGCCTCGTCT ACGTAGAGCA GGTAAAGGT CTGTCCCCGA 15480
ATGCTCTGCA GACACGGAAA GACACAAAAG AGGGGCTCAT AAGGGGCTAA CAGTAAAGGA 15540
GAGGAGGCGA ACAGTGCCTG GCTCTTGGTT CTTGGGAATA AAAGGGGGCG TGTGTGCCGA 15600
TCGATCGTAT GGGTGAGCCA GTGGATCCTG GACATGTGGT GAATGAGAAA GATTTTGAGG 15660
AGTGTGAACA ATTTTTCAGT CAACCCCTTA GGGAGCAAGT GGTCGCGGGG GTCAGGGCAC 15720
TCGACGGCCT CGGTCTCGCT GACTCTCTAT GTCACAAAAC AGAAAGACTC TGCCTGCTGA 15780
TGGACCTGGT GGGCACGGAG TGCTTTGCCA GGGTGTGCCG CCTAGACACC GGTGCGAAAT 15840
GAAGAGTGTG GCGAGTCCCT TATGTCAGTT CCACGGCGTG TTTTGCCTGT ACCAGTGTCTG 15900
CCAGTGCTG GCATACCACG TGTGTGATGG GGGCGCCGAA TGCGTTCTCC TGCATACGCC 15960
GGAGAGCGTC ATCTGCGAAC TAACGGGTAA CTGCATGCTC GGCAACATTC AAGAGGGCCA 16020
GTTTTTAGGG CCGGTACCGT ATCGGACTTT GGATAACCAG GTTGACAGGG ACGCATATCA 16080
CGGGATGCTA GCGTGTCTGA AACGGGACAT TGTGCGGTAT TTGCAGACAT GGCCGACAC 16140
CACCGTAATC GTGCAGGAAA TAGCCCTGGG GGACGGCGTC ACCGACACCA TCTCGGCCAT 16200
TATAGATGAA ACATTCCGTG AGTGTCTTCC CGTACTGGGG GAGGCCCAAG GCGGGTACCC 16260
CCTGGTCTGT AGCATGTATC TGCACGTTAT CGTCTCCATC TATTCGACAA AAACGGTGTA 16320
CAACAGTATG CTATTTAAAT GCACAAAGAA TAAAAAGTAC GACTGCATTG CCAAGCGGGT 16380
GCGGACAAAA TGGATGCGCA TGCTATCAAC GAAAGATACG TAGGTCTCTG CTGCCACCGT 16440
TTGGCCCACG TGGTGTCTGCC TAGGACCTTT CTGCTGCATC ACGCCATACC CCTGGAGCCC 16500
GAGATCATCT TTTCCACCTA CACCCGGTTC AGCCGGTCGC CAGGGTCATC CCGCCGGTTG 16560

006627479.066000

GTGGTGTGTG GGAAACGTGT CCTGCCAGGG GAGGAAACC AACTTGCCTC TTCACCTTCT	16620
GGTTTGGCGC TTAGCCTGCC TCTGTTTTCC CACGATGGGA ACTTTTCATCC ATTTGACATC	16680
TCGGTACTGC GCATTTCTGT CCCTGGTTCT AATCTTAGTC TTACTGTGAG ATTTCTCTAT	16740
CTATCTCTGG TGGTGGCTAT GGGGGCGGGA CGGAATAATG CGCGGAGTCC GACCGTTGAC	16800
GGGGTATCGC CGCCAGAGGG CGCCGTAGCC CACCCTTTGG AGGAACTGCA GAGGCTGGCG	16860
CGTGCTACGC CGGACCCGGC ACTCACCCTG GGACCGTTGC AGGTCTGTGAC CGGCCTTCTC	16920
CGCGCAGGGT CAGACGGAGA CCGCGCCACT CACCACATGG CGCTCGAGGC TCCGGGAACC	16980
GTGCGTGGAG AAAGCCTAGA CCCGCTGTGTT TCACAGAAGG GGCCAGCGCG CACACGCCAC	17040
AGGCCACCCC CCGTGCGACT GAGCTTCAAC CCCGTCAATG CCGATGTACC CGCTACCTGG	17100
CGAGACGCCA CTAACGTGTA CTCGGGTGCT CCCTACTATG TGTGTGTTTA CGAACGCGGT	17160
GGCCGTCAGG AAGACGACTG GCTGCCGATA CCACTGAGCT TCCAGAAGA GCCCGTGCCC	17220
CCGCCACCGG GCTTAGTGTT CATGGACGAC TTGTTTATTA ACACGAAGCA GTGCGACTTT	17280
GTGGACACGC TAGAGGCCGC CTGTGCGACG CAAGGCTACA CGTTGAGACA GCGCGTGCCT	17340
GTCGCCATTC CTCGCGACGC GGAAATCGCA GACGAGTTA AATCGCACTT TTTAGAGGCG	17400
TGCCTAGTGT TACGGGGGCT GGCTTCGGAG GCTAGTCCCT GGATAAGAGC TGCCACGTCC	17460
CCGCCCCCTG GCCGCCACGC CTGCTGGATG GACGTGTTAG GATTATGGGA AAGCCGCCCC	17520
CACACTCTAG GTTTGGAGTT ACGCGGCGTA AACTGTGGCG GCACGGACGG TGA CTGGTTA	17580
GAGATTTTAA AACAGCCCCG TGTGCAAAAG ACAGTCAGCG GGAGTCTTGT GGCATGCGTG	17640
ATCGTCACAC CCGCATTGGA AGCCTGGCTT GTGTTACCTG GGGGTTTTGC TATTAAAGCC	17700
CGCTATAGGG CGTCGAAGGA GGATCTGGTG TTCATTGAG GCGGCTATGG CTAGCCGGAG	17760
GCGCAAACCT CGGAATTTCC TAAACAAGGA ATGCATATGG ACTGTTAACC CAATGTCAGG	17820
GGACCATATC AAGGTCTTTA ACGCCTGCAC CTCTATCTCG CCGGTGTATG ACCCTGAGCT	17880
GGTAACCAGC TACGCACTGA GCGTGCCCTGC TTACAATGTG TCTGTGGCTA TCTTGCTGCA	17940
TAAAGTCATG GGACCGTGTG TGGCTGTGGG AATTAACGGA GAAATGATCA TGTACGTCGT	18000
AAGCCAGTGT GTTTCTGTGC GGCCCGTCCC GGGGCGCGAT GGTATGGCGC TCATCTACTT	18060
TGGACAGTTT CTGGAGGAAG CATCCGGA CTGAGATTCCC TACATTGCTC CCGCGCCGTC	18120
GCGCGAACAC GTACCTGACC TGACCAGACA AGAATTAGTT CATACTCCC AGGTGGTGCG	18180
CCGCGGCGAC CTGACCAATT GCACTATGGG TCTCGAATTC AGGAATGTGA ACCCTTTTGT	18240
TTGGCTCGGG GCGGATCGG TGTGGCTGCT GTTCTTGGGC GTGGACTACA TGGCGTCTG	18300
TCCGGGTGTC GACGGAATGC CGTCGTTGGC AAGAGTGGCC GCCCTGCTTA CCAGGTGCGA	18360
CCACCCAGAC TGTGTCCACT GCCATGGA CTGCTGGACAC GTTAATGTAT TTCGTGGGTA	18420
CTGTTCTGCG CAGTCGCCGG GTCTATCTAA CATCTGTCCC TGTATCAAAT CATGTGGGAC	18480
CGGGAATGGA GTGACTAGGG TCACTGGAAA CAGAAATTTT CTGGGTCTTC GTTTCGATCC	18540
CATTGTCCAG AGCAGGGTAA CAGCTCTGAA GATAACTAGC CACCCAACCC CCACGCACGT	18600

CGAGAATGTG	CTAACAGGAG	TGCTCGACGA	CGGCACCTTG	GTGCCGTCCG	TCCAAGGCAC	18660
CCTGGGTCCT	CTTACGAATG	TCTGACTACT	TCAGCCGCTT	GCTGATATAT	GAGTGTAATA	18720
AACTTAAGGC	CCTGGGCTTA	CGTTCTTATT	GAAGCATGTT	GCGCACATCA	GCGAGCTGGA	18780
CCGTCCTCCG	GGTCGCGTGT	AGATTATGGT	TCCGTTCTCC	TTCTTGATGT	TTAAATTTT	18840
GGGGGGGAAC	CACCGACAAA	SCGTCCTTTAT	GATTTCCGCG	AACACGGAGT	TGGCTACGTG	18900
CTTTTGGTGG	GCTACGTACC	CAATGTTAAT	GTTCTCTACG	GATGCCAGTA	GCATGCTGAT	18960
GATCGCCACC	ACTATCCATG	TCTTTCCGTG	TCTCCTTGGT	ATTAGGAATA	CGCTTGCCTT	19020
TTGCTTAAAC	GTCTGTAAAA	CACGTTTGG	AGTTTCAAAT	AAACCGAAGT	ACTGCTTAAA	19080
CAATCCAAAC	AACTGGTGCG	TCTTTTGTGG	GGCCTTGATT	GAAACCAAAA	AGAAAAAAGT	19140
GTGCATTACT	AGCTGCTGTT	GGAAGGGCTC	CAGCCAGTGC	ACCCCGGGAA	CGTAACAGCC	19200
GTTCAGAAAG	GACGAAAGGT	TAACCAGAAA	AGCCTGAAGT	TCGCGGTAGA	CAGAGCAGGC	19260
GTGCAGGGAG	TCGTGTGTTT	TTCTGCCCGC	CTGGTACTCG	ACCAGTTGAT	CGGCCGTGGA	19320
GACGTGCGCG	TCCTCGCGCA	CACACCGCAT	CTGCAAGTAT	GTTGATAGGG	ACTCCAATAG	19380
GCGCGGCTTT	GCGGGGACGT	TGTCCTCGGA	CGGTCTGGGG	GTTCCACAGT	CGGGATTGTC	19440
TGACGTGGGC	GTGGCGGGAT	GGTGCCGTGT	GCAGTATGTT	TCCAGGACCG	AACTGTATGA	19500
GTTTATTCTG	TGCACCACGC	CAATAAAAGG	GTGCGCCATC	CGTGCCGTTT	TGGGACAGTG	19560
TCGCGTGAAT	GTCGGGGCAC	TCAGTTCCCA	CCTCTCTCCG	GCGTCTTTGG	CGGTCTCCTC	19620
CAGGTTGGCG	GCAAGGCGCT	CCCTGTGACG	GCTGAGCAGC	ATGTTTGCTT	TGAGCTCGCT	19680
CGTGTCCGAG	GGTGACCCGG	AGGTGACCAG	TAGGTACGTC	AAGGGCGTAC	AACTTGCCCT	19740
GGACCTTAGC	GAGAACACAC	CTGGACAATT	TAAGTTGATA	GAAACTCCCC	TGAACAGCTT	19800
CCTCTTG GTT	TCCAACGTGA	TGCCCCGAGG	CCAGCCAATC	TGCAGTGGCC	GGCCGGCCTT	19860
GCGGCCAGAC	TTTAGTAATC	TCCACTTGCC	TAGACTGGAG	AAGCTCCAGA	GAGTCTCGG	19920
GCAGGGTTTC	GGGGCGGCGG	GTGAGGAAAT	CGCACTGGAC	CCGTCTCAGC	TAGAAACACA	19980
CGAAAAGGGC	CAGGTGTTCT	ACAACCACTA	TGCTACCGAG	GAGTGGACGT	GGGCTTTGAC	20040
TCTGAATAAG	GATGCGCTCC	TCGGGAGGC	TGTAGATGGC	CTGTGTGACG	CCGGAACTTG	20100
GAAGGGTCTT	CTTCCTGACG	ACCCCTTCC	GTTGCTATGG	CTGCTGTTCA	ACGGACCCGC	20160
CTCTTTTTGT	CGGGCCGACT	GTTGCCTGTA	CAAGCAGCAC	TGCGGTTACC	CGGGCCCGGT	20220
GCTACTTCCA	GGTCACATGT	ACGCTCCCAA	ACGGGATCTT	TTGTCGTTCC	TTAATCATGC	20280
CCTGAAGTAC	ACCAAGTTTC	TATACGGAGA	TTTTTCCGGG	ACATGGGCGG	CGGCTTGCCG	20340
CCCGCCATTG	GCTACTTCTC	GGATACAAAG	GGTAGTGAGT	CAGATGAAAA	TCATAGATGC	20400
TTCCGACACT	TACATTTCCT	ACACCTGCCT	CTTGTGTCAC	ATATATCAGC	AAAATAGCAT	20460
AATTGCGGGT	CAGGGGACCC	ACGTGGGTGG	AATCCTACTG	TTGAGTGGAA	AAGGGACCCA	20520
GTATATAACA	GGCAATGTTC	AGACCCAAAG	GTGTCCAAC	ACGGGCGACT	ATCTAATCAT	20580
CCCATCGTAT	GACATACCGG	CGATCATCAC	CATGATCAAG	GAGAATGGAC	TCAACCAACT	20640

188

CTAAAAGAGA GTTTATTAAG TCGGCTCTGG AGGCCAACAT CAACAGGAGG GCAGCTGTAT 20700
CGCTATTTGA 20710

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4131 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..4131
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

ATG GAG GCG ACC TTG GAG CAA CGA CTT TTC CCG TAC CTC GCC ACG GAG 48
Met Glu Ala Thr Leu Glu Gln Arg Pro Phe Pro Tyr Leu Ala Thr Glu
1 5 10 15

GCC AAC CTC CTA ACG CAG ATT AAG GAG TCG GCT GCC GAC GGA CTC TTC 96
Ala Asn Leu Leu Thr Gln Ile Lys Glu Ser Ala Ala Asp Gly Leu Phe
20 25 30

AAG AGC TTT CAG CTA TTG CTC GGC AAG GAC GCC AGA GAA GGC AGT GTC 144
Lys Ser Phe Gln Leu Leu Leu Gly Lys Asp Ala Arg Glu Gly Ser Val
35 40 45

CGT TTC GAA GCG CTA CTG GGC GTA TAT ACC AAT GTG GTG GAG TTT GTT 192
Arg Phe Glu Ala Leu Leu Gly Val Tyr Thr Asn Val Val Glu Phe Val
50 55 60

AAG TTT CTG GAG ACC GCC CTC GCC GCC GCT TGC GTC AAT ACC GAG TTC 240
Lys Phe Leu Glu Thr Ala Leu Ala Ala Ala Cys Val Asn Thr Glu Phe
65 70 75 80

AAG GAC CTG CGG AGA ATG ATA GAT GGA AAA ATA CAG TTT AAA ATT TCA 288
Lys Asp Leu Arg Arg Met Ile Asp Gly Lys Ile Gln Phe Lys Ile Ser
85 90 95

ATG CCC ACT ATT GCC CAC GGA GAC GGG AGG AGG CCC AAC AAG CAG AGA 336
Met Pro Thr Ile Ala His Gly Asp Gly Arg Arg Pro Asn Lys Gln Arg
100 105 110

CAG TAT ATC GTC ATG AAG GCT TGC AAT AAG CAC CAC ATC GGT GCG GAG 384
Gln Tyr Ile Val Met Lys Ala Cys Asn Lys His His Ile Gly Ala Glu
115 120 125

ATT GAG CTT GCG GCC GCA GAC ATC GAG CTT CTC TTC GCC GAG AAA GAG 432
Ile Glu Leu Ala Ala Ala Asp Ile Glu Leu Leu Phe Ala Glu Lys Glu
130 135 140

ACG CCC TTG GAC TTC ACA GAG TAC GCG GGT GCC ATC AAG ACG ATT ACG 480
Thr Pro Leu Asp Phe Thr Glu Tyr Ala Gly Ala Ile Lys Thr Ile Thr
145 150 155 160

006690" 64E20350

189

TCG	GCT	TTG	CAG	TTT	GGT	ATG	GAC	GCC	CTA	GAA	CGG	GGG	CTA	GTG	GAC	528
Ser	Ala	Leu	Gln	Phe	Gly	Met	Asp	Ala	Leu	Glu	Arg	Gly	Leu	Val	Asp	
			165						170					175		
ACG	GTT	CTC	GCA	GTT	AAA	CTT	CGG	CAC	GCT	CCA	CCC	GTC	TTT	ATT	TTA	576
Thr	Val	Leu	Ala	Val	Lys	Leu	Arg	His	Ala	Pro	Pro	Val	Phe	Ile	Leu	
			180					185					190			
AAG	ACG	CTG	GGC	GAT	CCC	GTC	TAC	TCT	GAG	AGG	GGC	CTC	AAA	AAG	GCC	624
Lys	Thr	Leu	Gly	Asp	Pro	Val	Tyr	Ser	Glu	Arg	Gly	Leu	Lys	Lys	Ala	
		195				200						205				
GTC	AAG	TCT	GAC	ATG	GTA	TCC	ATG	TTC	AAG	GCA	CAC	CTC	ATA	GAA	CAT	672
Val	Lys	Ser	Asp	Met	Val	Ser	Met	Phe	Lys	Ala	His	Leu	Ile	Glu	His	
	210					215						220				
TCA	TTT	TTT	CTA	GAT	AAG	GCC	GAG	CTC	ATG	ACA	AGG	GGG	AAG	CAG	TAT	720
Ser	Phe	Phe	Leu	Asp	Lys	Ala	Glu	Leu	Met	Thr	Arg	Gly	Lys	Gln	Tyr	
	225				230					235					240	
GTC	CTA	ACC	ATG	CTC	TCC	GAC	ATG	CTG	GCC	GCG	GTG	TGC	GAG	GAT	ACC	768
Val	Leu	Thr	Met	Leu	Ser	Asp	Met	Leu	Ala	Ala	Val	Cys	Glu	Asp	Thr	
			245					250						255		
GTC	TTT	AAG	GGT	GTC	AGC	ACG	TAC	ACC	ACG	GCC	TCT	GGG	CAG	CAG	GTG	816
Val	Phe	Lys	Gly	Val	Ser	Thr	Tyr	Thr	Thr	Ala	Ser	Gly	Gln	Gln	Val	
			260					265					270			
GCC	GGC	GTC	CTG	GAG	ACG	ACG	GAC	AGC	GTC	ATG	AGA	CGG	CTG	ATG	AAC	864
Ala	Gly	Val	Leu	Glu	Thr	Thr	Asp	Ser	Val	Met	Arg	Arg	Leu	Met	Asn	
		275					280					285				
CTG	CTG	GGG	CAA	GTG	GAA	AGT	GCC	ATG	TCC	GGG	CCC	GCG	GCC	TAC	GCC	912
Leu	Leu	Gly	Gln	Val	Glu	Ser	Ala	Met	Ser	Gly	Pro	Ala	Ala	Tyr	Ala	
		290				295					300					
AGC	TAC	GTT	GTC	AGG	GGT	GCC	AAC	CTC	GTC	ACC	GCC	GTT	AGC	TAC	GGA	960
Ser	Tyr	Val	Val	Arg	Gly	Ala	Asn	Leu	Val	Thr	Ala	Val	Ser	Tyr	Gly	
	305				310					315					320	
AGG	GCG	ATG	AGA	AAC	TTT	GAA	CAG	TTT	ATG	GCA	CGC	ATA	GTG	GAC	CAT	1008
Arg	Ala	Met	Arg	Asn	Phe	Glu	Gln	Phe	Met	Ala	Arg	Ile	Val	Asp	His	
			325						330					335		
CCC	AAC	GCT	CTG	CCG	TCT	GTG	GAA	GGT	GAC	AAG	GCC	GCT	CTG	GCG	GAC	1056
Pro	Asn	Ala	Leu	Pro	Ser	Val	Glu	Gly	Asp	Lys	Ala	Ala	Leu	Ala	Asp	
			340					345					350			
GGA	CAC	GAC	GAG	ATT	CAG	AGA	ACC	CGC	ATC	GCC	GCC	TCT	CTC	GTC	AAG	1104
Gly	His	Asp	Glu	Ile	Gln	Arg	Thr	Arg	Ile	Ala	Ala	Ser	Leu	Val	Lys	
		355					360					365				
ATA	GGG	GAT	AAG	TTT	GTG	GCC	ATT	GAA	AGT	TTG	CAG	CGC	ATG	TAC	AAC	1152
Ile	Gly	Asp	Lys	Phe	Val	Ala	Ile	Glu	Ser	Leu	Gln	Arg	Met	Tyr	Asn	
	370					375					380					
GAG	ACT	CAG	TTT	CCC	TGC	CCA	CTG	AAC	CGG	CGC	ATC	CAG	TAC	ACC	TAT	1200
Glu	Thr	Gln	Phe	Pro	Cys	Pro	Leu	Asn	Arg	Arg	Ile	Gln	Tyr	Thr	Tyr	
	385				390					395					400	
TTC	TTC	CCT	GTT	GGC	CTT	CAC	CTT	CCC	GTG	CCC	CGC	TAC	TCG	ACA	TCC	1248
Phe	Phe	Pro	Val	Gly	Leu	His	Leu	Pro	Val	Pro	Arg	Tyr	Ser	Thr	Ser	
			405						410					415		
GTC	TCA	GTC	AGG	GGC	GTA	GAA	TCC	CCG	GCC	ATC	CAG	TCG	ACC	GAG	ACG	1296
Val	Ser	Val	Arg	Gly	Val	Glu	Ser	Pro	Ala	Ile	Gln	Ser	Thr	Glu	Thr	
			420					425					430			

00507479.063900

[illegible]

TGG	GTG	GTT	AAT	AAA	AAC	AAC	GTG	CCT	CCT	TGC	TTC	GGT	TAC	CAA	AAC	1344
Trp	Val	Val	Asn	Lys	Asn	Asn	Val	Pro	Leu	Cys	Phe	Gly	Tyr	Gln	Asn	
		435					440					445				
GCC	CTC	AAA	AGC	ATA	TGC	CAC	CCT	CGA	ATG	CAC	AAC	CCC	ACC	CAG	TCA	1392
Ala	Leu	Lys	Ser	Ile	Cys	His	Pro	Arg	Met	His	Asn	Pro	Thr	Gln	Ser	
		450				455					460					
GCC	CAG	GCA	CTA	AAC	CAA	GCT	TTT	CCC	GAT	CCC	GAC	GGG	GGA	CAT	GGG	1440
Ala	Gln	Ala	Leu	Asn	Gln	Ala	Phe	Pro	Asp	Pro	Asp	Gly	Gly	His	Gly	
465					470					475					480	
TAC	GGT	CTC	AGG	TAT	GAG	CAG	ACG	CCA	AAC	ATG	AAC	CTA	TTC	AGA	ACG	1488
Tyr	Gly	Leu	Arg	Tyr	Glu	Gln	Thr	Pro	Asn	Met	Asn	Leu	Phe	Arg	Thr	
				485					490					495		
TTC	CAC	CAG	TAT	TAC	ATG	GGG	AAA	AAC	GTG	GCA	TTT	GTT	CCC	GAT	GTG	1536
Phe	His	Gln	Tyr	Tyr	Met	Gly	Lys	Asn	Val	Ala	Phe	Val	Pro	Asp	Val	
			500					505					510			
GCC	CAA	AAA	GCG	CTC	GTA	ACC	ACG	GAG	GAT	CTA	CTG	CAC	CCA	ACC	TCT	1584
Ala	Gln	Lys	Ala	Leu	Val	Thr	Thr	Glu	Asp	Leu	Leu	His	Pro	Thr	Ser	
		515					520					525				
CAC	CGT	CTC	CTC	AGA	TTG	GAG	GTC	CAC	CCC	TTC	TTT	GAT	TTT	TTT	GTG	1632
His	Arg	Leu	Leu	Arg	Leu	Glu	Val	His	Pro	Phe	Phe	Asp	Phe	Phe	Val	
	530					535					540					
CAC	CCC	TGT	CCT	GGA	GCG	AGA	GGA	TCG	TAC	CGC	GCC	ACC	CAC	AGA	ACA	1680
His	Pro	Cys	Pro	Gly	Ala	Arg	Gly	Ser	Tyr	Arg	Ala	Thr	His	Arg	Thr	
545				550						555					560	
ATG	GTT	GGA	AAT	ATA	CCA	CAA	CCG	CTC	GCT	CCA	AGG	GAG	TTT	CAG	GAA	1728
Met	Val	Gly	Asn	Ile	Pro	Gln	Pro	Leu	Ala	Pro	Arg	Glu	Phe	Gln	Glu	
				565					570					575		
AGT	AGA	GGG	GCG	CAG	TTC	GAC	GCT	GTG	ACG	AAT	ATG	ACA	CAC	GTC	ATA	1776
Ser	Arg	Gly	Ala	Gln	Phe	Asp	Ala	Val	Thr	Asn	Met	Thr	His	Val	Ile	
			580					585					590			
GAC	CAG	CTA	ACT	ATT	GAC	GTC	ATA	CAG	GAG	ACG	GCA	TTT	GAC	CCC	GCG	1824
Asp	Gln	Leu	Thr	Ile	Asp	Val	Ile	Gln	Glu	Thr	Ala	Phe	Asp	Pro	Ala	
		595					600					605				
TAT	CCC	CTG	TTC	TGC	TAT	GTA	ATC	GAA	GCA	ATG	ATT	CAC	GGA	CAG	GAA	1872
Tyr	Pro	Leu	Phe	Cys	Tyr	Val	Ile	Glu	Ala	Met	Ile	His	Gly	Gln	Glu	
	610					615					620					
GAA	AAA	TTC	GTG	ATG	AAC	ATG	CCC	CTC	ATT	GCC	CTG	GTC	ATT	CAA	ACC	1920
Glu	Lys	Phe	Val	Met	Asn	Met	Pro	Leu	Ile	Ala	Leu	Val	Ile	Gln	Thr	
625				630						635					640	
TAC	TGG	GTC	AAC													

ATC ACA CAT CTG GTT TCG GCT CTC CTC GAC CCG CAT CTG CTG CCT CCC Ile Thr His Leu Val Ser Ala Leu Leu Asp Pro His Leu Leu Pro Pro 705 710 715 720	2160
TTT GCC TAC CAC GAT GTC TTT ACG GAT CTT ATG CAG AAG TCA TCC AGA Phe Ala Tyr His Asp Val Phe Thr Asp Leu Met Gln Lys Ser Ser Arg 725 730 735	2208
CAA CCC ATA ATC AAG ATC GGG GAT CAA AAC TAC GAC AAC CCT CAA AAT Gln Pro Ile Ile Lys Ile Gly Asp Gln Asn Tyr Asp Asn Pro Gln Asn 740 745 750	2256
AGG GCG ACA TTC ATC AAC CTC AGG GGT CGC ATG GAG GAC CTA GTC AAT Arg Ala Thr Phe Ile Asn Leu Arg Gly Arg Met Glu Asp Leu Val Asn 755 760 765	2304
AAC CTT GTT AAC ATT TAC CAG ACA AGG GTC AAT GAG GAC CAT GAC GAG Asn Leu Val Asn Ile Tyr Gln Thr Arg Val Asn Glu Asp His Asp Glu 770 775 780	2352
AGA CAC GTC CTG GAC GTG GCG CCC CTG GAC GAG AAT GAC TAC AAC CCG Arg His Val Leu Asp Val Ala Pro Leu Asp Glu Asn Asp Tyr Asn Pro 785 790 795 800	2400
GTC CTC GAG AAG CTA TTC TAC TAT GTT TTA ATG CCG GTG TGC AGT AAC Val Leu Glu Lys Leu Phe Tyr Tyr Val Leu Met Pro Val Cys Ser Asn 805 810 815	2448
GGC CAC ATG TGC GGT ATG GGG GTC GAC TAT CAA AAC GTG GCC CTG ACG Gly His Met Cys Gly Met Gly Val Asp Tyr Gln Asn Val Ala Leu Thr 820 825 830	2496
CTG ACT TAC AAC GGC CCC GTC TTT GCG GAC GTC GTG AAC GCA CAG GAT Leu Thr Tyr Asn Gly Pro Val Phe Ala Asp Val Val Asn Ala Gln Asp 835 840 845	2544
GAT ATT CTA CTG CAC CTG GAG AAC GGA ACC TTG AAG GAC ATT CTG CAG Asp Ile Leu Leu His Leu Glu Asn Gly Thr Leu Lys Asp Ile Leu Gln 850 855 860	2592
GCA GGC GAC ATA CGC CCG ACG GTG GAC ATG ATC AGG GTG CTG TGC ACC Ala Gly Asp Ile Arg Pro Thr Val Asp Met Ile Arg Val Leu Cys Thr 865 870 875 880	2640
TCG TTT CTG ACG TGC CCT TTC GTC ACC CAG GCC GCT CGC GTG ATC ACA Ser Phe Leu Thr Cys Pro Phe Val Thr Gln Ala Ala Arg Val Ile Thr 885 890 895	2688
AAG CGG GAC CCG GCC CAG AGT TTT GCC ACG CAC GAA TAC GGG AAG GAT Lys Arg Asp Pro Ala Gln Ser Phe Ala Thr His Glu Tyr Gly Lys Asp 900 905 910	2736
GTG GCG CAG ACC GTG CTT GTT AAT GGC TTT GGT GCG TTC GCG GTG GCG Val Ala Gln Thr Val Leu Val Asn Gly Phe Gly Ala Phe Ala Val Ala 915 920 925	2784
GAC CGC TCT CGC GAG GCG GCG GAG ACT ATG TTT TAT CCG GTA CCC TTT Asp Arg Ser Arg Glu Ala Ala Glu Thr Met Phe Tyr Pro Val Pro Phe 930 935 940	2832
AAC AAG CTC TAC GCT GAC CCG TTG GTG GCT GCC ACA CTG CAT CCG CTC Asn Lys Leu Tyr Ala Asp Pro Leu Val Ala Ala Thr Leu His Pro Leu 945 950 955 960	2880
CTG CCA AAC TAT GTC ACC AGG CTC CCC AAC CAG AGA AAC GCG GTG GTC Leu Pro Asn Tyr Val Thr Arg Leu Pro Asn Gln Arg Asn Ala Val Val 965 970 975	2928

[illegible]

TTT	AAC	GTG	CCA	TCC	AAT	CTC	ATG	GCA	GAA	TAT	GAG	GAA	TGG	CAC	AAG	2976
Phe	Asn	Val	Pro	Ser	Asn	Leu	Met	Ala	Glu	Tyr	Glu	Glu	Trp	His	Lys	
		980						985					990			
TCG	CCC	GTC	GCG	GCG	TAT	GCC	GCG	TCT	TGT	CAG	GCC	ACC	CCG	GGC	GCC	3024
Ser	Pro	Val	Ala	Ala	Tyr	Ala	Ala	Ser	Cys	Gln	Ala	Thr	Pro	Gly	Ala	
		995					1000					1005				
ATT	AGC	GCC	ATG	GTG	AGC	ATG	CAC	CAA	AAA	CTA	TCT	GCC	CCC	AGT	TTC	3072
Ile	Ser	Ala	Met	Val	Ser	Met	His	Gln	Lys	Leu	Ser	Ala	Pro	Ser	Phe	
	1010					1015				1020						
ATT	TGC	CAG	GCA	AAA	CAC	CGC	ATG	CAC	CCT	GGT	TTT	GCC	ATG	ACA	GTC	3120
Ile	Cys	Gln	Ala	Lys	His	Arg	Met	His	Pro	Gly	Phe	Ala	Met	Thr	Val	
	1025				1030					1035					1040	
GTC	AGG	ACG	GAC	GAG	GTT	CTA	GCA	GAG	CAC	ATC	CTA	TAC	TGC	TCC	AGG	3168
Val	Arg	Thr	Asp	Glu	Val	Leu	Ala	Glu	His	Ile	Leu	Tyr	Cys	Ser	Arg	
			1045					1050						1055		
GCG	TCG	ACA	TCC	ATG	TTT	GTG	GGC	TTG	CCT	TCG	GTG	GTA	CGG	CGC	GAG	3216
Ala	Ser	Thr	Ser	Met	Phe	Val	Gly	Leu	Pro	Ser	Val	Val	Arg	Arg	Glu	
			1060				1065						1070			
GTA	CGT	TCG	GAC	GCG	GTG	ACT	TTT	GAA	ATT	ACC	CAC	GAG	ATC	GCT	TCC	3264
Val	Arg	Ser	Asp	Ala	Val	Thr	Phe	Glu	Ile	Thr	His	Glu	Ile	Ala	Ser	
	1075					1080				1085						
CTG	CAC	ACC	GCA	CTT	GGC	TAC	TCA	TCA	GTC	ATC	GCC	CCG	GCC	CAC	GTG	3312
Leu	His	Thr	Ala	Leu	Gly	Tyr	Ser	Ser	Val	Ile	Ala	Pro	Ala	His	Val	
	1090					1095				1100						
GCC	GCC	ATA	ACT	ACA	GAC	ATG	GGA	GTA	CAT	TGT	CAG	GAC	CTC	TTT	ATG	3360
Ala	Ala	Ile	Thr	Thr	Asp	Met	Gly	Val	His	Cys	Gln	Asp	Leu	Phe	Met	
	1105				1110					1115					1120	
ATT	TTC	CCA	GGG	GAC	GCG	TAT	CAG	GAC	CGC	CAG	CTG	CAT	GAC	TAT	ATC	3408
Ile	Phe	Pro	Gly	Asp	Ala	Tyr	Gln	Asp	Arg	Gln	Leu	His	Asp	Tyr	Ile	
			1125					1130						1135		
AAA	ATG	AAA	GCG	GGC	GTG	CAA	ACC	GGC	TCA	CCG	GGA	AAC	AGA	ATG	GAT	3456
Lys	Met	Lys	Ala	Gly	Val	Gln	Thr	Gly	Ser	Pro	Gly	Asn	Arg	Met	Asp	
		1140						1145					1150			
CAC	GTG	GGA	TAC	ACT	GCT	GGG	GTT	CCT	CGC	TGC	GAG	AAC	CTG	CCC	GGT	3504
His	Val	Gly	Tyr	Thr	Ala	Gly	Val	Pro	Arg	Cys	Glu	Asn	Leu	Pro	Gly	
	1155					1160						1165				
TTG	AGT	CAT	GGT	CAG	CTG	GCA	ACC	TGC	GAG	ATA	ATT	CCC	ACG	CCG	GTC	3552
Leu	Ser	His	Gly	Gln	Leu	Ala	Thr	Cys	Glu	Ile	Ile	Pro	Thr	Pro	Val	
	1170				1175					1180			</			

193

CTA TAC AAT ATC ACC TTT CGC CAG ACT GCG CTG CCG GGC ATG TAC AGT 3792
 Leu Tyr Asn Ile Thr Phe Arg Gln Thr Ala Leu Pro Gly Met Tyr Ser
 1250 1255 1260

CCT TGT CGG CAG TTC TTC CAC AAG GAA GAC ATT ATG CGG TAC AAT AGG 3840
 Pro Cys Arg Gln Phe Phe His Lys Glu Asp Ile Met Arg Tyr Asn Arg
 1265 1270 1275 1280

GGG TTG TAC ACT TTG GTT AAT GAG TAT TCT GCC AGG CTT GCT GGG GCC 3888
 Gly Leu Tyr Thr Leu Val Asn Glu Tyr Ser Ala Arg Leu Ala Gly Ala
 1285 1290 1295

CCC GCC ACC AGC ACT ACA GAC CTC CAG TAC GTC GTG GTC AAC GGT ACA 3936
 Pro Ala Thr Ser Thr Thr Asp Leu Gln Tyr Val Val Val Asn Gly Thr
 1300 1305 1310

GAC GTG TTT TTG GAC CAG CCT TGC CAT ATG CTG CAG GAG GCC TAT CCC 3984
 Asp Val Phe Leu Asp Gln Pro Cys His Met Leu Gln Glu Ala Tyr Pro
 1315 1320 1325

ACG CTC GCC GCC AGC CAC AGA GTT ATG CTT GCC GAG TAC ATG TCA AAC 4032
 Thr Leu Ala Ala Ser His Arg Val Met Leu Ala Glu Tyr Met Ser Asn
 1330 1335 1340

AAG CAG ACA CAC GCC CCA GTA CAC ATG GGC CAG TAT CTC ATT GAA GAG 4080
 Lys Gln Thr His Ala Pro Val His Met Gly Gln Tyr Leu Ile Glu Glu
 1345 1350 1355 1360

GTG GCG CCG ATG AAG AGA CTA TTA AAG CTC GGA AAC AAG GTG GTG TAT 4128
 Val Ala Pro Met Lys Arg Leu Leu Lys Leu Gly Asn Lys Val Val Tyr
 1365 1370 1375

TAG 4131

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1376 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Met Glu Ala Thr Leu Glu Gln Arg Pro Phe Pro Tyr Leu Ala Thr Glu
 1 5 10 15

Ala Asn Leu Leu Thr Gln Ile Lys Glu Ser Ala Ala Asp Gly Leu Phe
 20 25 30

Lys Ser Phe Gln Leu Leu Leu Gly Lys Asp Ala Arg Glu Gly Ser Val
 35 40 45

Arg Phe Glu Ala Leu Leu Gly Val Tyr Thr Asn Val Val Glu Phe Val
 50 55 60

Lys Phe Leu Glu Thr Ala Leu Ala Ala Ala Cys Val Asn Thr Glu Phe
 65 70 75 80

Lys Asp Leu Arg Arg Met Ile Asp Gly Lys Ile Gln Phe Lys Ile Ser
 85 90 95

Met Pro Thr Ile Ala His Gly Asp Gly Arg Arg Pro Asn Lys Glu Arg
 100 105 110

194

Gln Tyr Ile Val Met Lys Ala Cys Asn Lys His His Ile Gly Ala Glu
115 120 125
Ile Glu Leu Ala Ala Ala Asp Ile Glu Leu Leu Phe Ala Glu Lys Glu
130 135 140
Thr Pro Leu Asp Phe Thr Glu Tyr Ala Gly Ala Ile Lys Thr Ile Thr
145 150 155 160
Ser Ala Leu Gln Phe Gly Met Asp Ala Leu Glu Arg Gly Leu Val Asp
165 170 175
Thr Val Leu Ala Val Lys Leu Arg His Ala Pro Pro Val Phe Ile Leu
180 185 190
Lys Thr Leu Gly Asp Pro Val Tyr Ser Glu Arg Gly Leu Lys Lys Ala
195 200 205
Val Lys Ser Asp Met Val Ser Met Phe Lys Ala His Leu Ile Glu His
210 215 220
Ser Phe Phe Leu Asp Lys Ala Glu Leu Met Thr Arg Gly Lys Gln Tyr
225 230 235 240
Val Leu Thr Met Leu Ser Asp Met Leu Ala Ala Val Cys Glu Asp Thr
245 250 255
Val Phe Lys Gly Val Ser Thr Tyr Thr Thr Ala Ser Gly Gln Gln Val
260 265 270
Ala Gly Val Leu Glu Thr Thr Asp Ser Val Met Arg Arg Leu Met Asn
275 280 285
Leu Leu Gly Gln Val Glu Ser Ala Met Ser Gly Pro Ala Ala Tyr Ala
290 295 300
Ser Tyr Val Val Arg Gly Ala Asn Leu Val Thr Ala Val Ser Tyr Gly
305 310 315 320
Arg Ala Met Arg Asn Phe Glu Gln Phe Met Ala Arg Ile Val Asp His
325 330 335
Pro Asn Ala Leu Pro Ser Val Glu Gly Asp Lys Ala Ala Leu Ala Asp
340 345 350
Gly His Asp Glu Ile Gln Arg Thr Arg Ile Ala Ala Ser Leu Val Lys
355 360 365
Ile Gly Asp Lys Phe Val Ala Ile Glu Ser Leu Gln Arg Met Tyr Asn
370 375 380
Glu Thr Gln Phe Pro Cys Pro Leu Asn Arg Arg Ile Gln Tyr Thr Tyr
385 390 395 400
Phe Phe Pro Val Gly Leu His Leu Pro Val Pro Arg Tyr Ser Thr Ser
405 410 415
Val Ser Val Arg Gly Val Glu Ser Pro Ala Ile Gln Ser Thr Glu Thr
420 425 430
Trp Val Val Asn Lys Asn Asn Val Pro Leu Cys Phe Gly Tyr Gln Asn
435 440 445
Ala Leu Lys Ser Ile Cys His Pro Arg Met His Asn Pro Thr Gln Ser
450 455 460
Ala Gln Ala Leu Asn Gln Ala Phe Pro Asp Pro Asp Gly Gly His Gly

09507479.062900

480

465
Tyr Gly Leu Arg Tyr Glu Gln Thr Pro Asn Met Asn Leu Phe Arg Thr
485 490 495

Phe His Gln Tyr Tyr Met Gly Lys Asn Val Ala Phe Val Pro Asp Val
500 505 510

Ala Gln Lys Ala Leu Val Thr Thr Glu Asp Leu Leu His Pro Thr Ser
515 520 525

His Arg Leu Leu Arg Leu Glu Val His Pro Phe Phe Asp Phe Phe Val
530 535 540

His Pro Cys Pro Gly Ala Arg Gly Ser Tyr Arg Ala Thr His Arg Thr
545 550 555 560

Met Val Gly Asn Ile Pro Gln Pro Leu Ala Pro Arg Glu Phe Gln Glu
565 570 575

Ser Arg Gly Ala Gln Phe Asp Ala Val Thr Asn Met Thr His Val Ile
580 585 590

Asp Gln Leu Thr Ile Asp Val Ile Gln Glu Thr Ala Phe Asp Pro Ala
595 600 605

Tyr Pro Leu Phe Cys Tyr Val Ile Glu Ala Met Ile His Gly Gln Glu
610 615 620

Glu Lys Phe Val Met Asn Met Pro Leu Ile Ala Leu Val Ile Gln Thr
625 630 635 640

Tyr Trp Val Asn Ser Gly Lys Leu Ala Phe Val Asn Ser Tyr His Met
645 650 655

Val Arg Phe Ile Cys Thr His Ile Gly Asn Gly Ser Ile Pro Lys Glu
660 665 670

Ala His Gly His Tyr Arg Lys Ile Leu Gly Glu Leu Ile Ala Leu Glu
675 680 685

Gln Ala Leu Leu Lys Leu Ala Gly His Glu Thr Val Gly Arg Thr Pro
690 695 700

Ile Thr His Leu Val Ser Ala Leu Leu Asp Pro His Leu Leu Pro Pro
705 710 715 720

Phe Ala Tyr His Asp Val Phe Thr Asp Leu Met Gln Lys Ser Ser Arg
725 730 735

Gln Pro Ile Ile Lys Ile Gly Asp Gln Asn Tyr Asp Asn Pro Gln Asn
740 745 750

Arg Ala Thr Phe Ile Asn Leu Arg Gly Arg Met Glu Asp Leu Val Asn
755 760 765

Asn Leu Val Asn Ile Tyr Gln Thr Arg Val Asn Glu Asp His Asp Glu
770 775 780

Arg His Val Leu Asp Val Ala Pro Leu Asp Glu Asn Asp Tyr Asn Pro
785 790 795 800

Val Leu Glu Lys Leu Phe Tyr Tyr Val Leu Met Pro Val Cys Ser Asn
805 810 815

Gly His Met Cys Gly Met Gly Val Asp Tyr Gln Asn Val Ala Leu Thr
820 825 830

196

Leu Thr Tyr Asn Gly Pro Val Phe Ala Asp Val Val Asn Ala Gln Asp
 835 840 845
 Asp Ile Leu Leu His Leu Glu Asn Gly Thr Leu Lys Asp Ile Leu Gln
 850 855 860
 Ala Gly Asp Ile Arg Pro Thr Val Asp Met Ile Arg Val Leu Cys Thr
 865 870 875 880
 Ser Phe Leu Thr Cys Pro Phe Val Thr Gln Ala Ala Arg Val Ile Thr
 885 890 895
 Lys Arg Asp Pro Ala Gln Ser Phe Ala Thr His Glu Tyr Gly Lys Asp
 900 905 910
 Val Ala Gln Thr Val Leu Val Asn Gly Phe Gly Ala Phe Ala Val Ala
 915 920 925
 Asp Arg Ser Arg Glu Ala Ala Glu Thr Met Phe Tyr Pro Val Pro Phe
 930 935 940
 Asn Lys Leu Tyr Ala Asp Pro Leu Val Ala Ala Thr Leu His Pro Leu
 945 950 955 960
 Leu Pro Asn Tyr Val Thr Arg Leu Pro Asn Gln Arg Asn Ala Val Val
 965 970 975
 Phe Asn Val Pro Ser Asn Leu Met Ala Glu Tyr Glu Glu Trp His Lys
 980 985 990
 Ser Pro Val Ala Ala Tyr Ala Ala Ser Cys Gln Ala Thr Pro Gly Ala
 995 1000 1005
 Ile Ser Ala Met Val Ser Met His Gln Lys Leu Ser Ala Pro Ser Phe
 1010 1015 1020
 Ile Cys Gln Ala Lys His Arg Met His Pro Gly Phe Ala Met Thr Val
 1025 1030 1035 1040
 Val Arg Thr Asp Glu Val Leu Ala Glu His Ile Leu Tyr Cys Ser Arg
 1045 1050 1055
 Ala Ser Thr Ser Met Phe Val Gly Leu Pro Ser Val Val Arg Arg Glu
 1060 1065 1070
 Val Arg Ser Asp Ala Val Thr Phe Glu Ile Thr His Glu Ile Ala Ser
 1075 1080 1085
 Leu His Thr Ala Leu Gly Tyr Ser Ser Val Ile Ala Pro Ala His Val
 1090 1095 1100
 Ala Ala Ile Thr Thr Asp Met Gly Val His Cys Gln Asp Leu Phe Met
 1105 1110 1115 1120
 Ile Phe Pro Gly Asp Ala Tyr Gln Asp Arg Gln Leu His Asp Tyr Ile
 1125 1130 1135
 Lys Met Lys Ala Gly Val Gln Thr Gly Ser Pro Gly Asn Arg Met Asp
 1140 1145 1150
 His Val Gly Tyr Thr Ala Gly Val Pro Arg Cys Glu Asn Leu Pro Gly
 1155 1160 1165
 Leu Ser His Gly Gln Leu Ala Thr Cys Glu Ile Ile Pro Thr Pro Val
 1170 1175 1180
 Thr Ser Asp Val Ala Tyr Phe Gln Thr Pro Ser Asn Pro Arg Gly Arg

006230 624963

197

1185 1190 1195 1200
 Ala Ala Ser Val Val Ser Cys Asp Ala Tyr Ser Asn Glu Ser Ala Glu
 1205 1210 1215
 Arg Leu Phe Tyr Asp His Ser Ile Pro Asp Pro Ala Tyr Glu Cys Arg
 1220 1225 1230
 Ser Thr Asn Asn Pro Trp Ala Ser Gln Arg Gly Ser Leu Gly Asp Val
 1235 1240 1245
 Leu Tyr Asn Ile Thr Phe Arg Gln Thr Ala Leu Pro Gly Met Tyr Ser
 1250 1255 1260
 Pro Cys Arg Gln Phe Phe His Lys Glu Asp Ile Met Arg Tyr Asn Arg
 1265 1270 1275 1280
 Gly Leu Tyr Thr Leu Val Asn Glu Tyr Ser Ala Arg Leu Ala Gly Ala
 1285 1290 1295
 Pro Ala Thr Ser Thr Thr Asp Leu Gln Tyr Val Val Val Asn Gly Thr
 1300 1305 1310
 Asp Val Phe Leu Asp Gln Pro Cys His Met Leu Gln Glu Ala Tyr Pro
 1315 1320 1325
 Thr Leu Ala Ala Ser His Arg Val Met Leu Ala Glu Tyr Met Ser Asn
 1330 1335 1340
 Lys Gln Thr His Ala Pro Val His Met Gly Gln Tyr Leu Ile Glu Glu
 1345 1350 1355 1360
 Val Ala Pro Met Lys Arg Leu Leu Lys Leu Gly Asn Lys Val Val Tyr
 1365 1370 1375

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1143 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..1143
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

AGC ATT CGG GGA CAG ACC TTT AAC CTG CTC TAC GTA GAC GAG GCG AAT 48
 Ser Ile Arg Gly Gln Thr Phe Asn Leu Leu Tyr Val Asp Glu Ala Asn
 1 5 10 15
 TTT ATT AAA AAG GAT GCA CTG CCG GCT ATT CTG GGT TTC ATG CTT CAG 96
 Phe Ile Lys Lys Asp Ala Leu Pro Ala Ile Leu Gly Phe Met Leu Gln
 20 25 30

AAA Lys	GAC Asp	GCC Ala	AAG Lys	CTT Leu	ATA Ile	TTT Phe	ATA Ile	TCA Ser	TCC Ser	GTG Val	AAC Asn	TCG Ser	TCA Ser	GAC Asp	CGC Arg	144
		35					40					45				
TCC Ser	ACG Thr	AGT Ser	TTC Phe	CTG Leu	CTT Leu	AAC Asn	CTC Leu	AGG Arg	AAC Asn	GCC Ala	CAG Gln	GAA Glu	AAG Lys	ATG Met	CTG Leu	192
	50					55					60					
AAT Asn	GTG Val	GTC Val	AGT Ser	TAC Tyr	GTG Val	TGT Cys	GCG Ala	GAC Asp	CAC His	CGA Arg	GAA Glu	GAT Asp	TTC Phe	CAC His	CTG Leu	240
	65				70					75					80	
CAA Gln	GAC Asp	GCA Ala	CTA Leu	GTG Val	TCC Ser	TGT Cys	CCT Pro	TGT Cys	TAC Tyr	AGA Arg	CTG Leu	CAC His	ATT Ile	CCG Pro	ACG Thr	288
				85					90					95		
TAC Tyr	ATC Ile	ACC Thr	ATC Ile	GAC Asp	GAA Glu	TCC Ser	ATC Ile	AAA Lys	ACC Thr	ACC Thr	ACC Thr	AAC Asn	CTC Leu	TTT Phe	ATG Met	336
			100					105						110		
GAG Glu	GGG Gly	GCA Ala	TTC Phe	GAC Asp	ACC Thr	GAA Glu	CTA Leu	ATG Met	GGC Gly	GAG Glu	GGA Gly	GCA Ala	GCG Ala	TCG Ser	TCA Ser	384
		115					120					125				
AAT Asn	GCT Ala	ACG Thr	CTT Leu	TAC Tyr	CGC Arg	GTG Val	GTG Val	GGT Gly	GAC Asp	GCA Ala	GCG Ala	CTG Leu	ACA Thr	CAG Gln	TTT Phe	432
	130					135					140					
GAC Asp	ATG Met	TGT Cys	CGG Arg	GTA Val	GAC Asp	ACC Thr	ACC Thr	GCC Ala	CAG Gln	GAG Glu	GTT Val	CAG Gln	AAG Lys	TGC Cys	CTT Leu	480
	145				150					155					160	
GGA Gly	AAA Lys	CAG Gln	CTG Leu	TTT Phe	GTT Val	TAC Tyr	ATC Ile	GAC Asp	CCC Pro	GCG Ala	TAT Tyr	ACG Thr	AAC Asn	AAC Asn	ACG Thr	528
				165					170					175		
GAG Glu	GCG Ala	TCC Ser	GGT Gly	ACT Thr	GGC Gly	GTG Val	GGC Gly	GCC Ala	GTT Val	GTC Val	ACG Thr	AGT Ser	ACT Thr	CAG Gln	ACT Thr	576
			180				185						190			
CCC Pro	ACC Thr	AGA Arg	AGC Ser	CTC Leu	ATA Ile	TTG Leu	GGC Gly	ATG Met	GAG Glu	CAT His	TTT Phe	TTT Phe	CTG Leu	CGC Arg	GAC Asp	624
		195					200					205				
CTC Leu	ACT Thr	GGC Gly	GCA Ala	GCT Ala	GCT Ala	TAC Tyr	GAG Glu	ATA Ile	GCG Ala	TCC Ser	TGC Cys	GCA Ala	TGC Cys	ACG Thr	ATG Met	672
	210					215					220					
ATT Ile	AAG Lys	GCG Ala	ATC Ile	GCT Ala	GTG Val	CTC Leu	CAC His	ACC Thr	ACA Thr	ATT Ile	GAG Glu	CGC Arg	GTG Val	AAC Asn	GCG Ala	720
	225				230				235					240		
GCG Ala	GTC Val	GAA Glu	GGC Gly	AAC Asn	AGC Ser	AGC Ser	CAA Gln	GAT Asp	TCT Ser	GGG Gly	GTG Val	GCC Ala	ATT Ile	GCA Ala	ACC Thr	768
				245				250						255		
GTC Val	CTT Leu	AAC Asn	GAA Glu	ATA Ile	TGC Cys	CCG Pro	CTC Leu	CCC Pro	ATA Ile	CAT His	TTT Phe	CTA Leu	CAC His	TAT Tyr	ACT Thr	816
			260					265					270			
GAC Asp	AAG Lys	AGC Ser	AGC Ser	GCC Ala	CTG Leu	CAG Gln	TGG Trp	CCA Pro	ATT Ile	TAC Tyr	ATG Met	TTG Leu	GGA Gly	GGC Gly	GAG Glu	864
		275														

199

CTG AGC GCC AGC CAG ACG GTG GTG TCC AAC ACC ATC AAA ATA TCA TTT	960
Leu Ser Ala Ser Gln Thr Val Val Ser Asn Thr Ile Lys Ile Ser Phe	
305 310 315 320	
GAC CCG GTG ACC TAC CTG GTA GAA CAG GTC CGC GCG ATC AAG TGC GTC	1008
Asp Pro Val Thr Tyr Leu Val Glu Gln Val Arg Ala Ile Lys Cys Val	
325 330 335	
CCG CTT AGG GAT GGA GGG CAG TCA TAC AGC GCC AAG CAA AAG CAC ATG	1056
Pro Leu Arg Asp Gly Gly Gln Ser Tyr Ser Ala Lys Gln Lys His Met	
340 345 350	
TCG GAC GAC TTA CTT GTG GCA GTT GTC ATG GCC CAT TTT ATG GCT ACC	1104
Ser Asp Asp Leu Leu Val Ala Val Val Met Ala His Phe Met Ala Thr	
355 360 365	
GAT GAT AGA CAC ATG TAC AAG CCC ATA TCC CCA CAA TAA	1143
Asp Asp Arg His Met Tyr Lys Pro Ile Ser Pro Gln	
370 375 380	

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 380 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Ser Ile Arg Gly Gln Thr Phe Asn Leu Leu Tyr Val Asp Glu Ala Asn	
1 5 10 15	
Phe Ile Lys Lys Asp Ala Leu Pro Ala Ile Leu Gly Phe Met Leu Gln	
20 25 30	
Lys Asp Ala Lys Leu Ile Phe Ile Ser Ser Val Asn Ser Ser Asp Arg	
35 40 45	
Ser Thr Ser Phe Leu Leu Asn Leu Arg Asn Ala Gln Glu Lys Met Leu	
50 55 60	
Asn Val Val Ser Tyr Val Cys Ala Asp His Arg Glu Asp Phe His Leu	
65 70 75 80	
Gln Asp Ala Leu Val Ser Cys Pro Cys Tyr Arg Leu His Ile Pro Thr	
85 90 95	
Tyr Ile Thr Ile Asp Glu Ser Ile Lys Thr Thr Thr Asn Leu Phe Met	
100 105 110	
Glu Gly Ala Phe Asp Thr Glu Leu Met Gly Glu Gly Ala Ala Ser Ser	
115 120 125	
Asn Ala Thr Leu Tyr Arg Val Val Gly Asp Ala Ala Leu Thr Gln Phe	
130 135 140	
Asp Met Cys Arg Val Asp Thr Thr Ala Gln Glu Val Gln Lys Cys Leu	
145 150 155 160	
Gly Lys Gln Leu Phe Val Tyr Ile Asp Pro Ala Tyr Thr Asn Asn Thr	
165 170 175	
Glu Ala Ser Gly Thr Gly Val Gly Ala Val Val Thr Ser Thr Gln Thr	
180 185 190	

006290.04E960

Pro	Thr	Arg	Ser	Leu	Ile	Leu	Gly	Met	Glu	His	Phe	Leu	Arg	Asp
		195					200					205		
Leu	Thr	Gly	Ala	Ala	Ala	Tyr	Glu	Ile	Ala	Ser	Cys	Ala	Cys	Thr
	210					215					220			Met
Ile	Lys	Ala	Ile	Ala	Val	Leu	His	Thr	Thr	Ile	Glu	Arg	Val	Asn
225					230					235				240
Ala	Val	Glu	Gly	Asn	Ser	Ser	Gln	Asp	Ser	Gly	Val	Ala	Ile	Ala
				245					250					255
Val	Leu	Asn	Glu	Ile	Cys	Pro	Leu	Pro	Ile	His	Phe	Leu	His	Tyr
			260					265					270	Thr
Asp	Lys	Ser	Ser	Ala	Leu	Gln	Trp	Pro	Ile	Tyr	Met	Leu	Gly	Gly
		275					280					285		Glu
Lys	Ser	Ser	Ala	Phe	Glu	Thr	Phe	Ile	Tyr	Ala	Leu	Asn	Ser	Gly
	290					295					300			Thr
Leu	Ser	Ala	Ser	Gln	Thr	Val	Val	Ser	Asn	Thr	Ile	Lys	Ile	Ser
305					310					315				Phe
Asp	Pro	Val	Thr	Tyr	Leu	Val	Glu	Gln	Val	Arg	Ala	Ile	Lys	Cys
				325					330					335
Pro	Leu	Arg	Asp	Gly	Gly	Gln	Ser	Tyr	Ser	Ala	Lys	Gln	Lys	His
			340					345					350	Met
Ser	Asp	Asp	Leu	Leu	Val	Ala	Val	Val	Met	Ala	His	Phe	Met	Ala
		355					360					365		Thr
Asp	Asp	Arg	His	Met	Tyr	Lys	Pro	Ile	Ser	Pro	Gln			
	370					375					380			

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 234 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION: 1..234
(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

ATG	GGT	GAG	CCA	GTG	GAT	CCT	GGA	CAT	GTG	GTG	AAT	GAG	AAA	GAT	TTT
Met	Gly	Glu	Pro	Val	Asp	Pro	Gly	His	Val	Val	Asn	Glu	Lys	Asp	Phe
1				5					10					15	
GAG	GAG	TGT	GAA	CAA	TTT	TTC	AGT	CAA	CCC	CTT	AGG	GAG	CAA	GTG	GTC
Glu	Glu	Cys	Glu	Gln	Phe	Phe	Ser	Gln	Pro	Leu	Arg	Glu	Gln	Val	Val
			20					25					30		

48

96

201

GCG GGG GTC AGG GCA CTC GAC GGC CTC GGT CTC GCT GAC TCT CTA TGT	144
Ala Gly Val Arg Ala Leu Asp Gly Leu Gly Leu Ala Asp Ser Leu Cys	
35 40 45	
CAC AAA ACA GAA AGA CTC TGC CTG CTG ATG GAC CTG GTG GGC ACG GAG	192
His Lys Thr Glu Arg Leu Cys Leu Leu Met Asp Leu Val Gly Thr Glu	
50 55 60	
TGC TTT GCG AGG GTG TGC CGC CTA GAC ACC GGT GCG AAA TGA	234
Cys Phe Ala Arg Val Cys Arg Leu Asp Thr Gly Ala Lys	
65 70 75	

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 77 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Met Gly Glu Pro Val Asp Pro Gly His Val Val Asn Glu Lys Asp Phe	
1 5 10 15	
Glu Glu Cys Glu Gln Phe Phe Ser Gln Pro Leu Arg Glu Gln Val Val	
20 25 30	
Ala Gly Val Arg Ala Leu Asp Gly Leu Gly Leu Ala Asp Ser Leu Cys	
35 40 45	
His Lys Thr Glu Arg Leu Cys Leu Leu Met Asp Leu Val Gly Thr Glu	
50 55 60	
Cys Phe Ala Arg Val Cys Arg Leu Asp Thr Gly Ala Lys	
65 70 75	

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 585 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..585
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

ATG AAG AGT GTG GCG AGT CCC TTA TGT CAG TTC CAC GGC GTG TTT TGC	48
Met Lys Ser Val Ala Ser Pro Leu Cys Gln Phe His Gly Val Phe Cys	
1 5 10 15	
CTG TAC CAG TGT CGC CAG TGC CTG GCA TAC CAC GTG TGT GAT GGG GGC	96

005290"64740960

202

Leu	Tyr	Gln	Cys	Arg	Gln	Cys	Leu	Ala	Tyr	His	Val	Cys	Asp	Gly	Gly		
			20							25				30			
GCC	GAA	TGC	GTT	CTC	CTG	CAT	ACG	CCG	GAG	AGC	GTC	ATC	TGC	GAA	CTA	144	
Ala	Glu	Cys	Val	Leu	Leu	His	Thr	Pro	Glu	Ser	Val	Ile	Cys	Glu	Leu		
			35				40				45						
ACG	GGT	AAC	TGC	ATG	CTC	GGC	AAC	ATT	CAA	GAG	GGC	CAG	TTT	TTA	GGG	192	
Thr	Gly	Asn	Cys	Met	Leu	Gly	Asn	Ile	Gln	Glu	Gly	Gln	Phe	Leu	Gly		
			50				55				60						
CCG	GTA	CCG	TAT	CGG	ACT	TTG	GAT	AAC	CAG	GTT	GAC	AGG	GAC	GCA	TAT	240	
Pro	Val	Pro	Tyr	Arg	Thr	Leu	Asp	Asn	Gln	Val	Asp	Arg	Asp	Ala	Tyr		
			65				70				75				80		
CAC	GGG	ATG	CTA	GCG	TGT	CTG	AAA	CGG	GAC	ATT	GTG	CGG	TAT	TTG	CAG	288	
His	Gly	Met	Leu	Ala	Cys	Leu	Lys	Arg	Asp	Ile	Val	Arg	Tyr	Leu	Gln		
				85				90				95					
ACA	TGG	CCG	GAC	ACC	ACC	GTA	ATC	GTG	CAG	GAA	ATA	GCC	CTG	GGG	GAC	336	
Thr	Trp	Pro	Asp	Thr	Thr	Val	Ile	Val	Gln	Glu	Ile	Ala	Leu	Gly	Asp		
			100				105				110						
GGC	GTC	ACC	GAC	ACC	ATC	TCG	GCC	ATT	ATA	GAT	GAA	ACA	TTC	GGT	GAG	384	
Gly	Val	Thr	Asp	Thr	Ile	Ser	Ala	Ile	Ile	Asp	Glu	Thr	Phe	Gly	Glu		
			115				120				125						
TGT	CTT	CCC	GTA	CTG	GGG	GAG	GCC	CAA	GGC	GGG	TAC	GCC	CTG	GTC	TGT	432	
Cys	Leu	Pro	Val	Leu	Gly	Glu	Ala	Gln	Gly	Gly	Tyr	Ala	Leu	Val	Cys		
			130				135				140						
AGC	ATG	TAT	CTG	CAC	GTT	ATC	GTC	TCC	ATC	TAT	TCG	ACA	AAA	ACG	GTG	480	
Ser	Met	Tyr	Leu	His	Val	Ile	Val	Ser	Ile	Tyr	Ser	Thr	Lys	Thr	Val		
			145				150				155				160		
TAC	AAC	AGT	ATG	CTA	TTT	AAA	TGC	ACA	AAG	AAT	AAA	AAG	TAC	GAC	TGC	528	
Tyr	Asn	Ser	Met	Leu	Phe	Lys	Cys	Thr	Lys	Asn	Lys	Lys	Tyr	Asp	Cys		
			165				170				175						
ATT	GCC	AAG	CGG	GTG	CGG	ACA	AAA	TGG	ATG	CGC	ATG	CTA	TCA	ACG	AAA	576	
Ile	Ala	Lys	Arg	Val	Arg	Thr	Lys	Trp	Met	Arg	Met	Leu	Ser	Thr	Lys		
			180				185				190						
GAT	ACG	TAG														585	
Asp	Thr	.															
			195														

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 194 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Met Lys Ser Val Ala Ser Pro Leu Cys Gln Phe His Gly Val Phe Cys
1 5 10 15
Leu Tyr Gln Cys Arg Gln Cys Leu Ala Tyr His Val Cys Asp Gly Gly
20 25 30
Ala Glu Cys Val Leu Leu His Thr Pro Glu Ser Val Ile Cys Glu Leu
35 40 45

203

Thr Gly Asn Cys Met Leu Gly Asn Ile Gln Glu Gly Gln Phe Leu Gly
 50 55 60
 Pro Val Pro Tyr Arg Thr Leu Asp Asn Gln Val Asp Arg Asp Ala Tyr
 65 70 75 80
 His Gly Met Leu Ala Cys Leu Lys Arg Asp Ile Val Arg Tyr Leu Gln
 85 90 95
 Thr Trp Pro Asp Thr Thr Val Ile Val Gln Glu Ile Ala Leu Gly Asp
 100 105 110
 Gly Val Thr Asp Thr Ile Ser Ala Ile Ile Asp Glu Thr Phe Gly Glu
 115 120 125
 Cys Leu Pro Val Leu Gly Glu Ala Gln Gly Gly Tyr Ala Leu Val Cys
 130 135 140
 Ser Met Tyr Leu His Val Ile Val Ser Ile Tyr Ser Thr Lys Thr Val
 145 150 155 160
 Tyr Asn Ser Met Leu Phe Lys Cys Thr Lys Asn Lys Lys Tyr Asp Cys
 165 170 175
 Ile Ala Lys Arg Val Arg Thr Lys Trp Met Arg Met Leu Ser Thr Lys
 180 185 190
 Asp Thr

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 939 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: N
- (iv) ANTI-SENSE: N
- (ix) FEATURE:
 - (A) NAME/KEY: CDS
 - (B) LOCATION: 1..939
 - (D) OTHER INFORMATION:
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

ATG GCT AGC CGG AGG CGC AAA CTT CGG AAT TTC CTA AAC AAG GAA TGC	48
Met Ala Ser Arg Arg Arg Lys Leu Arg Asn Phe Leu Asn Lys Glu Cys	
1 5 10 15	
ATA TGG ACT GTT AAC CCA ATG TCA GGG GAC CAT ATC AAG GTC TTT AAC	96
Ile Trp Thr Val Asn Pro Met Ser Gly Asp His Ile Lys Val Phe Asn	
20 25 30	
GCC TGC ACC TCT ATC TCG CCG GTG TAT GAC CCT GAG CTG GTA ACC AGC	144
Ala Cys Thr Ser Ile Ser Pro Val Tyr Asp Pro Glu Leu Val Thr Ser	
35 40 45	
TAC GCA CTG AGC GTG CCT GCT TAC AAT GTG TCT GTG GCT ATC TTG CTG	192
Tyr Ala Leu Ser Val Pro Ala Tyr Asn Val Ser Val Ala Ile Leu Leu	
50 55 60	

006230-642360

204

CAT	AAA	GTC	ATG	GGA	CCG	TGT	GTG	GCT	GTG	GGA	ATT	AAC	GGA	GAA	ATG	240
His	Lys	Val	Met	Gly	Pro	Cys	Val	Ala	Val	Gly	Ile	Asn	Gly	Glu	Met	
65					70					75					80	
ATC	ATG	TAC	GTC	GTA	AGC	CAG	TGT	GTT	TCT	GTG	CGG	CCC	GTC	CCG	GGG	288
Ile	Met	Tyr	Val	Val	Ser	Gln	Cys	Val	Ser	Val	Arg	Pro	Val	Pro	Gly	
			85						90					95		
CGC	GAT	GGT	ATG	GCG	CTC	ATC	TAC	TTT	GGA	CAG	TTT	CTG	GAG	GAA	GCA	336
Arg	Asp	Gly	Met	Ala	Leu	Ile	Tyr	Phe	Gly	Gln	Phe	Leu	Glu	Glu	Ala	
			100					105					110			
TCC	GGA	CTG	AGA	TTT	CCC	TAC	ATT	GCT	CCG	CCG	CCG	TCG	CGC	GAA	CAC	384
Ser	Gly	Leu	Arg	Phe	Pro	Tyr	Ile	Ala	Pro	Pro	Pro	Ser	Arg	Glu	His	
		115					120					125				
GTA	CCT	GAC	CTG	ACC	AGA	CAA	GAA	TTA	GTT	CAT	ACC	TCC	CAG	GTG	GTG	432
Val	Pro	Asp	Leu	Thr	Arg	Gln	Glu	Leu	Val	His	Thr	Ser	Gln	Val	Val	
	130					135					140					
CGC	CGC	GGC	GAC	CTG	ACC	AAT	TGC	ACT	ATG	GGT	CTC	GAA	TTC	AGG	AAT	480
Arg	Arg	Gly	Asp	Leu	Thr	Asn	Cys	Thr	Met	Gly	Leu	Glu	Phe	Arg	Asn	
145				150						155					160	
GTG	AAC	CCT	TTT	GTT	TGG	CTC	GGG	GGC	GGA	TCG	GTG	TGG	CTG	CTG	TTC	528
Val	Asn	Pro	Phe	Val	Trp	Leu	Gly	Gly	Gly	Ser	Val	Trp	Leu	Leu	Phe	
				165					170					175		
TTG	GGC	GTG	GAC	TAC	ATG	GCG	TTC	TGT	CCG	GGT	GTC	GAC	GGA	ATG	CCG	576
Leu	Gly	Val	Asp	Tyr	Met	Ala	Phe	Cys	Pro	Gly	Val	Asp	Gly	Met	Pro	
			180					185					190			
TCG	TTG	GCA	AGA	GTG	GCC	GCC	CTG	CTT	ACC	AGG	TGC	GAC	CAC	CCA	GAC	624
Ser	Leu	Ala	Arg	Val	Ala	Ala	Leu	Leu	Thr	Arg	Cys	Asp	His	Pro	Asp	
		195					200					205				
TGT	GTC	CAC	TGC	CAT	GGA	CTC	CGT	GGA	CAC	GTT	AAT	GTA	TTT	CGT	GGG	672
Cys	Val	His	Cys	His	Gly	Leu	Arg	Gly	His	Val	Asn	Val	Phe	Arg	Gly	
	210					215					220					
TAC	TGT	TCT	GCG	CAG	TCG	CCG	GGT	CTA	TCT	AAC	ATC	TGT	CCC	TGT	ATC	720
Tyr	Cys	Ser	Ala	Gln	Ser	Pro	Gly	Leu	Ser	Asn	Ile	Cys	Pro	Cys	Ile	
225					230					235					240	
AAA	TCA	TGT	GGG	ACC	GGG	AAT	GGA	GTG	ACT	AGG	GTC	ACT	GGA	AAC	AGA	768
Lys	Ser	Cys	Gly	Thr	Gly	Asn	Gly	Val	Thr	Arg	Val	Thr	Gly	Asn	Arg	
				245					250					255		
AAT	TTT	CTG	GGT	CTT	CTG	TTC	GAT	CCC	ATT	GTC	CAG	AGC	AGG	GTA	ACA	816
Asn	Phe	Leu	Gly	Leu	Leu	Phe	Asp	Pro	Ile	Val	Gln	Ser	Arg	Val	Thr	
			260					265					270			
GCT	CTG	AAG	ATA	ACT	AGC	CAC	CCA	ACC	CCC	ACG	CAC	GTC	GAG	AAT	GTG	864
Ala	Leu	Lys	Ile	Thr	Ser	His	Pro	Thr	Pro	Thr	His	Val	Glu	Asn	Val	
		275					280					285				
CTA	ACA	GGA	GTG	CTC	GAC	GAC	GGC	ACC	TTG	GTG	CCG	TCC	GTC	CAA	GGC	912
Leu	Thr	Gly	Val	Leu	Asp	Asp	Gly	Thr	Leu	Val	Pro	Ser	Val	Gln	Gly	
		290				295					300					
ACC	CTG	GGT	CCT	CTT	ACG	AAT	GTC	TGA								939
Thr	Leu	Gly	Pro	Leu	Thr	Asn	Val									
305					310											

(2) INFORMATION FOR SEQ ID NO:11:

205

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 312 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Met Ala Ser Arg Arg Arg Lys Leu Arg Asn Phe Leu Asn Lys Glu Cys
 1 5 10 15
 Ile Trp Thr Val Asn Pro Met Ser Gly Asp His Ile Lys Val Phe Asn
 20 25 30
 Ala Cys Thr Ser Ile Ser Pro Val Tyr Asp Pro Glu Leu Val Thr Ser
 35 40 45
 Tyr Ala Leu Ser Val Pro Ala Tyr Asn Val Ser Val Ala Ile Leu Leu
 50 55 60
 His Lys Val Met Gly Pro Cys Val Ala Val Gly Ile Asn Gly Glu Met
 65 70 75 80
 Ile Met Tyr Val Val Ser Gln Cys Val Ser Val Arg Pro Val Pro Gly
 85 90 95
 Arg Asp Gly Met Ala Leu Ile Tyr Phe Gly Gln Phe Leu Glu Glu Ala
 100 105 110
 Ser Gly Leu Arg Phe Pro Tyr Ile Ala Pro Pro Pro Ser Arg Glu His
 115 120 125
 Val Pro Asp Leu Thr Arg Gln Glu Leu Val His Thr Ser Gln Val Val
 130 135 140
 Arg Arg Gly Asp Leu Thr Asn Cys Thr Met Gly Leu Glu Phe Arg Asn
 145 150 155 160
 Val Asn Pro Phe Val Trp Leu Gly Gly Gly Ser Val Trp Leu Leu Phe
 165 170 175
 Leu Gly Val Asp Tyr Met Ala Phe Cys Pro Gly Val Asp Gly Met Pro
 180 185 190
 Ser Leu Ala Arg Val Ala Ala Leu Leu Thr Arg Cys Asp His Pro Asp
 195 200 205
 Cys Val His Cys His Gly Leu Arg Gly His Val Asn Val Phe Arg Gly
 210 215 220
 Tyr Cys Ser Ala Gln Ser Pro Gly Leu Ser Asn Ile Cys Pro Cys Ile
 225 230 235 240
 Lys Ser Cys Gly Thr Gly Asn Gly Val Thr Arg Val Thr Gly Asn Arg
 245 250 255
 Asn Phe Leu Gly Leu Leu Phe Asp Pro Ile Val Gln Ser Arg Val Thr
 260 265 270
 Ala Leu Lys Ile Thr Ser His Pro Thr Pro Thr His Val Glu Asn Val
 275 280 285
 Leu Thr Gly Val Leu Asp Asp Gly Thr Leu Val Pro Ser Val Gln Gly
 290 295 300
 Thr Leu Gly Pro Leu Thr Asn Val

0060749.062400

305

310

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 86 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..86
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

ATG	GAC	TCA	ACC	AAC	TCT	AAA	AGA	GAG	TTT	ATT	AAG	TCG	GCT	CTG	GAG	48
Met	Asp	Ser	Thr	Asn	Ser	Lys	Arg	Glu	Phe	Ile	Lys	Ser	Ala	Leu	Glu	
1				5				10						15		
GCC	AAC	ATC	AAC	AGG	AGG	GCA	GCT	GTA	TCG	CTA	TTT	GA				86
Ala	Asn	Ile	Asn	Arg	Arg	Ala	Ala	Val	Ser	Leu	Phe					
			20					25								

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Met	Asp	Ser	Thr	Asn	Ser	Lys	Arg	Glu	Phe	Ile	Lys	Ser	Ala	Leu	Glu
1				5				10						15	
Ala	Asn	Ile	Asn	Arg	Arg	Ala	Ala	Val	Ser	Leu	Phe				
			20					25							

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1743 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: N

(ix) FEATURE:

(B) LOCATION: 1..1743

(D) OTHER INFORMATION:

ATG GCA GAA GGC GGT TTT GGA GCG GAC TCG GTG GGG CGC GGC GGA GAA	48
Met Ala Glu Gly Gly Phe Gly Ala Asp Ser Val Gly Arg Gly Gly Glu	
1 5 10 15	
AAG GCC TCT GTG ACT AGG GGA GGC AGG TGG GAC TTG GGG AGC TCG GAC	96
Lys Ala Ser Val Thr Arg Gly Gly Arg Trp Asp Leu Gly Ser Ser Asp	
20 25 30	
GAC GAA TCA AGC ACC TCC ACA ACC AGC ACG GAT ATG GAC GAC CTC CCT	144
Asp Glu Ser Ser Thr Ser Thr Thr Ser Thr Asp Met Asp Asp Leu Pro	
35 40 45	
GAG GAG AGG AAA CCA CTA ACG GGA AAG TCT GTA AAA ACC TCG TAC ATA	192
Glu Glu Arg Lys Pro Leu Thr Gly Lys Ser Val Lys Thr Ser Tyr Ile	
50 55 60	
TAC GAC GTG CCC ACC GTC CCG ACC AGC AAG CCG TGG CAT TTA ATG CAC	240
Tyr Asp Val Pro Thr Val Pro Thr Ser Lys Pro Trp His Leu Met His	
65 70 75 80	
GAC AAC TCC CTC TAC GCA ACG CCT AGG TTT CCG CCC AGA CCT CTC ATA	288
Asp Asn Ser Leu Tyr Ala Thr Pro Arg Phe Pro Pro Arg Pro Leu Ile	
85 90 95	
CGG CAC CCT TCC GAA AAA GGC AGC ATT TTT GCC AGT CGG TTG TCA GCG	336
Arg His Pro Ser Glu Lys Gly Ser Ile Phe Ala Ser Arg Leu Ser Ala	
100 105 110	
ACT GAC GAC GAC TCG GGA GAC TAC GCG CCA ATG GAT CGC TTC GCC TTC	384
Thr Asp Asp Asp Ser Gly Asp Tyr Ala Pro Met Asp Arg Phe Ala Phe	
115 120 125	
CAG AGC CCC AGG GTG TGT GGT CGC CCT CCC CTT CCG CCT CCA AAT CAC	432
Gln Ser Pro Arg Val Cys Gly Arg Pro Pro Leu Pro Pro Pro Asn His	
130 135 140	
CCA CCT CCG GCA ACT AGG CCG GCA GAC GCG TCA ATG GGG GAC GTG GGC	480
Pro Pro Pro Ala Thr Arg Pro Ala Asp Ala Ser Met Gly Asp Val Gly	
145 150 155 160	
TGG GCG GAT CTG CAG GGA CTC AAG AGG ACC CCA AAG GGA TTT TTA AAA	528
Trp Ala Asp Leu Gln Gly Leu Lys Arg Thr Pro Lys Gly Phe Leu Lys	
165 170 175	
ACA TCT ACC AAG GGG GGC AGT CTC AAA GCC CGT GGA CGC GAT GTA GGT	576
Thr Ser Thr Lys Gly Gly Ser Leu Lys Ala Arg Gly Arg Asp Val Gly	
180 185 190	
GAC CGT CTC AGG GAC GGC GGC TTT GCC TTT AGT CCT AGG GGC GTG AAA	624
Asp Arg Leu Arg Asp Gly Gly Phe Ala Phe Ser Pro Arg Gly Val Lys	
195 200 205	
TCT GCC ATA GGG CAA AAC ATT AAA TCA TGG TTG GGG ATC GGA GAA TCA	672
Ser Ala Ile Gly Gln Asn Ile Lys Ser Trp Leu Gly Ile Gly Glu Ser	
210 215 220	
TCG GCG ACT GCT GTC CCC GTC ACC ACG CAG CTT ATG GTA CCG GTG CAC	720

Ser 225	Ala	Thr	Ala	Val	Pro 230	Val	Thr	Thr	Gln	Leu	Met	Val	Pro	Val	His 240	
CTC Leu	ATT Ile	AGA Arg	ACG Thr	CCT Pro	GTG Val	ACC Thr	GTG Val	GAC Asp	TAC Tyr	AGG Arg	AAT Asn	GTT Val	TAT Tyr	TTG Leu	CTT Leu	768
TAC Tyr	TTA Leu	GAG Glu	GGG Gly	GTA Val	ATG Met	GGT Gly	GTG Val	GGC Gly	AAA Lys	TCA Ser	ACG Thr	CTG Leu	GTC Val	AAC Asn	GCC Ala	816
GTG Val	TGC Cys	GGG Gly	ATC Ile	TTG Leu	CCC Pro	CAG Gln	GAG Glu	AGA Arg	GTG Val	ACA Thr	AGT Ser	TTT Phe	CCC Pro	GAG Glu	CCC Pro	864
ATG Met	GTG Val	TAC Tyr	TGG Trp	ACG Thr	AGG Arg	GCA Ala	TTT Phe	ACA Thr	GAT Asp	TGT Cys	TAC Tyr	AAG Lys	GAA Glu	ATT Ile	TCC Ser	912
CAC His 305	CTG Leu	ATG Met	AAG Lys	TCT Ser	GGT Gly	AAG Lys	GCG Ala	GGA Gly	GAC Asp	CCG Pro	CTG Leu	ACG Thr	TCT Ser	GCC Ala	AAA Lys 320	960
ATA Ile	TAC Tyr	TCA Ser	TGC Cys	CAA Gln	AAC Asn	AAG Lys	TTT Phe	TCG Ser	CTC Leu	CCC Pro	TTC Phe	CGG Arg	ACG Thr	AAC Asn	GCC Ala 335	1008
ACC Thr	GCT Ala	ATC Ile	CTG Leu	CGA Arg	ATG Met	ATG Met	CAG Gln	CCC Pro	TGG Trp	AAC Asn	GTT Val	GGG Gly	GGT Gly	GGG Gly	TCT Ser	1056
GGG Gly	AGG Arg	GGC Gly	ACT Thr	CAC His	TGG Trp	TGC Cys	GTC Val	TTT Phe	GAT Asp	AGG Arg	CAT His	CTC Leu	CTC Leu	TCC Ser	CCA Pro	1104
GCA Ala	GTG Val	GTG Val	TTC Phe	CCT Pro	CTC Leu	ATG Met	CAC His	CTG Leu	AAG Lys	CAC His	GGC Gly	CGC Arg	CTA Leu	TCT Ser	TTT Phe	1152
GAT Asp 385	CAC His	TTC Phe	TTT Phe	CAA Gln	TTA Leu	CTT Leu	TCC Ser	ATC Ile	TTT Phe	AGA Arg	GCC Ala	ACA Thr	GAA Glu	GGC Gly	GAC Asp 400	1200
GTG Val	GTC Val	GCC Ala	ATT Ile	CTC Leu	ACC Thr	CTC Leu	TCC Ser	AGC Ser	GCC Ala	GAG Glu	TCG Ser	TTG Leu	CGG Arg	CGG Arg	GTC Val	1248
AGG Arg	GCG Ala	AGG Arg	GGA Gly	AGA Arg	AAG Lys	AAC Asn	GAC Asp	GGG Gly	ACG Thr	GTG Val	GAG Glu	CAA Gln	AAC Asn	TAC Tyr	ATC Ile	1296
AGA Arg	GAA Glu	TTG Leu	GCG Ala	TGG Trp	GCT Ala	TAT Tyr	CAC His	GCC Ala	GTG Val	TAC Tyr	TGT Cys	TCA Ser	TGG Trp	ATC Ile	ATG Met	1344
TTG Leu	CAG Gln	TAC Tyr	ATC Ile	ACT Thr	GTG Val	GAG Glu	CAG Gln	ATG Met	GTA Val	CAA Gln	CTA Leu	TGC Cys	GTA Val	CAA Gln	ACC Thr	1392
ACA Thr 465	AAT Asn	ATT Ile	CCG Pro	GAA Glu	ATC Ile	TGC Cys	TTC Phe	CGC Arg	AGC Ser	GTG Val	CGC Arg	CTG Leu	GCA Ala	CAC His	AAG Lys 480	1440
GAG Glu	GAA Glu	ACT Thr	TTG Leu	AAA Lys	AAC Asn	CTT Leu	CAC His	GAG Glu	CAG Gln	AGC Ser	ATG Met	CTA Leu	CCT Pro	ATG Met	ATC Ile	1488

209

ACC	GGT	GTA	CTG	GAT	CCC	GTG	AGA	CAT	CAT	CCC	GTC	GTG	ATC	GAG	CTT	1536
Thr	Gly	Val	Leu	Asp	Pro	Val	Arg	His	His	Pro	Val	Val	Ile	Glu	Leu	
			500					505					510			
TGC	TTT	TGT	TTC	TTC	ACA	GAG	CTG	AGA	AAA	TTA	CAA	TTT	ATC	GTA	GCC	1564
Cys	Phe	Cys	Phe	Phe	Thr	Glu	Leu	Arg	Lys	Leu	Gln	Phe	Ile	Val	Ala	
		515					520					525				
GAC	GCG	GAT	AAG	TTC	CAC	GAC	GAC	GTA	TGC	GGC	CTG	TGG	ACC	GAA	ATC	1632
Asp	Ala	Asp	Lys	Phe	His	Asp	Asp	Val	Cys	Gly	Leu	Trp	Thr	Glu	Ile	
	530					535					540					
TAC	AGG	CAG	ATC	CTG	TCC	AAT	CCG	GCT	ATT	AAA	CCC	AGG	GCC	ATC	AAC	1680
Tyr	Arg	Gln	Ile	Leu	Ser	Asn	Pro	Ala	Ile	Lys	Pro	Arg	Ala	Ile	Asn	
545					550					555					560	
TGG	CCA	GCA	TTA	GAG	AGC	CAG	TCT	AAA	GCA	GTT	AAT	CAC	CTA	GAG	GAG	1728
Trp	Pro	Ala	Leu	Glu	Ser	Gln	Ser	Lys	Ala	Val	Asn	His	Leu	Glu	Glu	
			565						570					575		
ACA	TGC	AGG	GTC	TAG												1743
Thr	Cys	Arg	Val													
			580													

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 580 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Met	Ala	Glu	Gly	Gly	Phe	Gly	Ala	Asp	Ser	Val	Gly	Arg	Gly	Gly	Glu	
1				5					10					15		
Lys	Ala	Ser	Val	Thr	Arg	Gly	Gly	Arg	Trp	Asp	Leu	Gly	Ser	Ser	Asp	
			20					25					30			
Asp	Glu	Ser	Ser	Thr	Ser	Thr	Thr	Ser	Thr	Asp	Met	Asp	Asp	Leu	Pro	
		35					40					45				
Glu	Glu	Arg	Lys	Pro	Leu	Thr	Gly	Lys	Ser	Val	Lys	Thr	Ser	Tyr	Ile	
	50					55					60					
Tyr	Asp	Val	Pro	Thr	Val	Pro	Thr	Ser	Lys	Pro	Trp	His	Leu	Met	His	
65					70				75					80		
Asp	Asn	Ser	Leu	Tyr	Ala	Thr	Pro	Arg	Phe	Pro	Pro	Arg	Pro	Leu	Ile	
			85						90					95		
Arg	His	Pro	Ser	Glu	Lys	Gly	Ser	Ile	Phe	Ala	Ser	Arg	Leu	Ser	Ala	
		100						105					110			
Thr	Asp	Asp	Asp	Ser	Gly	Asp	Tyr	Ala	Pro	Met	Asp	Arg	Phe	Ala	Phe	
		115					120					125				
Gln	Ser	Pro	Arg	Val	Cys	Gly	Arg	Pro	Pro	Leu	Pro	Pro	Pro	Asn	His	
	130					135					140					
Pro	Pro	Pro	Ala	Thr	Arg	Pro	Ala	Asp	Ala	Ser	Met	Gly	Asp	Val	Gly	
145					150					155					160	
Trp	Ala	Asp	Leu	Gln	Gly	Leu	Lys	Arg	Thr	Pro	Lys	Gly	Phe	Leu	Lys	

165										170				175			
Thr	Ser	Thr	Lys 180	Gly	Gly	Ser	Leu	Lys 185	Ala	Arg	Gly	Arg	Asp 190	Val	Gly		
Asp	Arg	Leu 195	Arg	Asp	Gly	Gly	Phe 200	Ala	Phe	Ser	Pro	Arg 205	Gly	Val	Lys		
Ser	Ala 210	Ile	Gly	Gln	Asn 215	Ile	Lys	Ser	Trp	Leu	Gly 220	Ile	Gly	Glu	Ser		
Ser 225	Ala	Thr	Ala	Val	Pro 230	Val	Thr	Thr	Gln	Leu 235	Met	Val	Pro	Val	His 240		
Leu	Ile	Arg	Thr	Pro 245	Val	Thr	Val	Asp	Tyr 250	Arg	Asn	Val	Tyr	Leu 255	Leu		
Tyr	Leu	Glu	Gly 260	Val	Met	Gly	Val	Gly 265	Lys	Ser	Thr	Leu	Val	Asn	Ala		
Val	Cys	Gly 275	Ile	Leu	Pro	Gln	Glu 280	Arg	Val	Thr	Ser	Phe 285	Pro	Glu	Pro		
Met	Val 290	Tyr	Trp	Thr	Arg	Ala 295	Phe	Thr	Asp	Cys	Tyr 300	Lys	Glu	Ile	Ser		
His 305	Leu	Met	Lys	Ser	Gly 310	Lys	Ala	Gly	Asp	Pro 315	Leu	Thr	Ser	Ala	Lys 320		
Ile	Tyr	Ser	Cys	Gln 325	Asn	Lys	Phe	Ser	Leu 330	Pro	Phe	Arg	Thr	Asn 335	Ala		
Thr	Ala	Ile	Leu 340	Arg	Met	Met	Gln	Pro 345	Trp	Asn	Val	Gly	Gly 350	Gly	Ser		
Gly	Arg	Gly 355	Thr	His	Trp	Cys	Val 360	Phe	Asp	Arg	His	Leu 365	Leu	Ser	Pro		
Ala	Val 370	Val	Phe	Pro	Leu	Met 375	His	Leu	Lys	His	Gly 380	Arg	Leu	Ser	Phe		
Asp 385	His	Phe	Phe	Gln	Leu 390	Leu	Ser	Ile	Phe	Arg 395	Ala	Thr	Glu	Gly	Asp 400		
Val	Val	Ala	Ile	Leu 405	Thr	Leu	Ser	Ser	Ala 410	Glu	Ser	Leu	Arg	Arg 415	Val		
Arg	Ala	Arg	Gly 420	Arg	Lys	Asn	Asp	Gly 425	Thr	Val	Glu	Gln	Asn 430	Tyr	Ile		
Arg	Glu	Leu 435	Ala	Trp	Ala	Tyr	His 440	Ala	Val	Tyr	Cys	Ser 445	Trp	Ile	Met		
Leu	Gln 450	Tyr	Ile	Thr	Val	Glu 455	Gln	Met	Val	Gln	Leu 460	Cys	Val	Gln	Thr		
Thr 465	Asn	Ile	Pro	Glu	Ile 470	Cys	Phe	Arg	Ser	Val 475	Arg	Leu	Ala	His	Lys 480		
Glu	Glu	Thr	Leu	Lys 485	Asn	Leu	His	Glu	Gln 490	Ser	Met	Leu	Pro	Met 495	Ile		
Thr	Gly	Val	Leu 500	Asp	Pro	Val	Arg	His 505	His	Pro	Val	Val	Ile 510	Glu	Leu		
Cys	Phe	Cys 515	Phe	Phe	Thr	Glu	Leu 520	Arg	Lys	Leu	Gln	Phe 525	Ile	Val	Ala		

WO 96/06159

211

Asp Ala Asp Lys Phe His Asp Asp Val Cys Gly Leu Trp Thr Glu Ile
 530 535 540
 Tyr Arg Gln Ile Leu Ser Asn Pro Ala Ile Lys Pro Arg Ala Ile Asn
 545 550 555 560
 Trp Pro Ala Leu Glu Ser Gln Ser Lys Ala Val Asn His Leu Glu Glu
 565 570 575
 Thr Cys Arg Val
 580

(2) INFORMATION FOR SEQ ID NO:16:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 2193 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

- (ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION: 1..2193
 (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

ATG CAG GGT CTA GCC TTC TTG GCG GCC CTT GCA TGC TGG CGA TGC ATA	48
Met Gln Gly Leu Ala Phe Leu Ala Ala Leu Ala Cys Trp Arg Cys Ile	
1 5 10 15	
TCG TTG ACA TGT GGA GCC ACT GGC GCG TTG CCG ACA ACG GCG ACG ACA	96
Ser Leu Thr Cys Gly Ala Thr Gly Ala Leu Pro Thr Thr Ala Thr Thr	
20 25 30	
ATA ACC CGC TCC GCC ACG CAG CTC ATC AAT GGG AGA ACC AAC CTC TCC	144
Ile Thr Arg Ser Ala Thr Gln Leu Ile Asn Gly Arg Thr Asn Leu Ser	
35 40 45	
ATA GAA CTG GAA TTC AAC GGC ACT AGT TTT TTT CTA AAT TGG CAA AAT	192
Ile Glu Leu Glu Phe Asn Gly Thr Ser Phe Phe Leu Asn Trp Gln Asn	
50 55 60	
CTG TTG AAT GTG ATC ACG GAG CCG GCC CTG ACA GAG TTG TGG ACC TCC	240
Leu Leu Asn Val Ile Thr Glu Pro Ala Leu Thr Glu Leu Trp Thr Ser	
65 70 75 80	
GCC GAA GTC GCC GAG GAC CTC AGG GTA ACT CTG AAA AAG AGG CAA AGT	288
Ala Glu Val Ala Glu Asp Leu Arg Val Thr Leu Lys Lys Arg Gln Ser	
85 90 95	
CTT TTT TTC CCC AAC AAG ACA GTT GTG ATC TCT GGA GAC GGC CAT CGC	336
Leu Phe Phe Pro Asn Lys Thr Val Ile Ser Gly Asp Gly His Arg	
100 105 110	
TAT ACG TGC GAG GTG CCG ACG TCG TCG CAA ACT TAT AAC ATC ACC AAG	384
Tyr Thr Cys Glu Val Pro Thr Ser Ser Gln Thr Tyr Asn Ile Thr Lys	
115 120 125	
GGC TTT AAC TAT AGC GCT CTG CCC GGG CAC CTT GGC GGA TTT GGG ATC	432

006230-02120960

Gly 130	Phe 130	Asn	Tyr	Ser	Ala	Leu 135	Pro	Gly	His	Leu	Gly 140	Gly	Phe	Gly	Ile	
AAC Asn 145	GCG Ala	CGT Arg	CTG Leu	GTA Val	CTG Leu 150	GGT Gly	GAT Asp	ATC Ile	TTC Phe	GCA Ala 155	TCA Ser	AAA Lys	TGG Trp	TCG Ser	CTA Leu 160	480
TTC Phe	GCG Ala	AGG Arg	GAC Asp	ACC Thr 165	CCA Pro	GAG Glu	TAT Tyr	CGG Arg	GTG Val 170	TTT Phe	TAC Tyr	CCA Pro	ATG Met	AAT Asn 175	GTC Val	528
ATG Met	GCC Ala	GTC Val	AAG Lys 180	TTT Phe	TCC Ser	ATA Ile	TCC Ser	ATT Ile 185	GGC Gly	AAC Asn	AAC Asn	GAG Glu	TCC Ser 190	GGC Gly	GTA Val	576
GCG Ala	CTC Leu	TAT Tyr 195	GGA Gly	GTG Val	GTG Val	TCG Ser	GAA Glu 200	GAT Asp	TTC Phe	GTG Val	GTC Val	GTC Val 205	ACG Thr	CTC Leu	CAC His	624
AAC Asn 210	AGG Arg	TCC Ser	AAA Lys	GAG Glu	GCT Ala	AAC Asn 215	GAG Glu	ACG Thr	GCG Ala	TCC Ser	CAT His 220	CTT Leu	CTG Leu	TTC Phe	GGT Gly	672
CTC Leu 225	CCG Pro	GAT Asp	TCA Ser	CTG Leu	CCA Pro 230	TCT Ser	CTG Leu	AAG Lys	GGC Gly	CAT His 235	GCC Ala	ACC Thr	TAT Tyr	GAT Asp	GAA Glu 240	720
CTC Leu	ACG Thr	TTC Phe	GCC Ala	CGA Arg 245	AAC Asn	GCA Ala	AAA Lys	TAT Tyr	GCG Ala 250	CTA Leu	GTG Val	GCG Ala	ATC Ile	CTG Leu 255	CCT Pro	768
AAA Lys	GAT Asp	TCT Ser	TAC Tyr 260	CAG Gln	ACA Thr	CTC Leu	CTT Leu	ACA Thr 265	GAG Glu	AAT Asn	TAC Tyr	ACT Thr	CGC Arg 270	ATA Ile	TTT Phe	816
CTG Leu	AAC Asn	ATG Met 275	ACG Thr	GAG Glu	TCG Ser	ACG Thr	CCC Pro 280	CTC Leu	GAG Glu	TTC Phe	ACG Thr	CGG Arg 285	ACG Thr	ATC Ile	CAG Gln	864
ACC Thr	AGG Arg 290	ATC Ile	GTA Val	TCA Ser	ATC Ile	GAG Glu 295	GCC Ala	AGG Arg	CGC Arg	GCC Ala	TGC Cys 300	GCA Ala	GCT Ala	CAA Gln	GAG Glu	912
GCG Ala 305	GCG Ala	CCG Pro	GAC Asp	ATA Ile	TTC Phe 310	TTG Leu	GTG Val	TTG Leu	TTT Phe	CAG Gln 315	ATG Met	TTG Leu	GTG Val	GCA Ala	CAC His 320	960
TTT Phe	CTT Leu	GTT Val	GCG Ala	CGG Arg 325	GGC Gly	ATT Ile	GCC Ala	GAG Glu	CAC His 330	CGA Arg	TTT Phe	GTG Val	GAG Glu	GTG Val 335	GAC Asp	1008
TGC Cys	GTG Val	TGT Cys	CGG Arg 340	CAG Gln	TAT Tyr	GCG Ala	GAA Glu	CTG Leu 345	TAT Tyr	TTT Phe	CTC Leu	CGC Arg 350	CGC Arg	ATC Ile	TCG Ser	1056
CGT Arg	CTG Leu	TGC Cys 355	ATG Met	CCC Pro	ACG Thr	TTC Phe	ACC Thr 360	ACT Thr	GTC Val	GGG Gly	TAT Tyr	AAC Asn 365	CAC His	ACC Thr	ACC Thr	1104
CTT Leu	GGC Gly 370	GCT Ala	GTG Val	GCC Ala	GCC Ala	ACA Thr 375	CAA Gln	ATA Ile	GCT Ala	CGC Arg	GTG Val 380	TCC Ser	GCC Ala	ACG Thr	AAG Lys	1152
TTG Leu 385	GCC Ala	AGT Ser	TTG Leu	CCC Pro	CGC Arg 390	TCT Ser	TCC Ser	CAG Gln	GAA Glu	ACA Thr 395	GTG Val	CTG Leu	GCC Ala	ATG Met	GTC Val 400	1200

CAG Gln	CTT Leu	GGC Gly	GCC Ala	CGT Arg 405	GAT Asp	GGC Gly	GCC Ala	GTC Val	CCT Pro 410	TCC Ser	TCC Ser	ATT Ile	CTG Leu	GAG Glu 415	GGC Gly	1248
ATT Ile	GCT Ala	ATG Met	GTC Val 420	GTC Val	GAA Glu	CAT His	ATG Met	TAT Tyr 425	ACC Thr	GCC Ala	TAC Tyr	ACT Thr	TAT Tyr 430	GTG Val	TAC Tyr	1296
ACA Thr	CTC Leu	GGC Gly 435	GAT Asp	ACT Thr	GAA Glu	AGA Arg	AAA Lys 440	TTA Leu	ATG Met	TTG Leu	GAC Asp	ATA Ile 445	CAC His	ACG Thr	GTC Val	1344
CTC Leu	ACC Thr 450	GAC Asp	AGC Ser	TGC Cys	CCG Pro	CCC Pro 455	AAA Lys	GAC Asp	TCC Ser	GGA Gly	GTA Val 460	TCA Ser	GAA Glu	AAG Lys	CTA Leu	1392
CTG Leu 465	AGA Arg	ACA Thr	TAT Tyr	TTG Leu	ATG Met 470	TTC Phe	ACA Thr	TCA Ser	ATG Met	TGT Cys 475	ACC Thr	AAC Asn	ATA Ile	GAG Glu	CTG Leu 480	1440
GGC Gly	GAA Glu	ATG Met	ATC Ile 485	GCC Ala	CGC Arg	TTT Phe	TCC Ser	AAA Lys	CCG Pro 490	GAC Asp	AGC Ser	CTT Leu	AAC Asn	ATC Ile 495	TAT Tyr	1488
AGG Arg	GCA Ala	TTC Phe	TCC Ser 500	CCC Pro	TGC Cys	TTT Phe	CTA Leu	GGA Gly 505	CTA Leu	AGG Arg	TAC Tyr	GAT Asp	TTG Leu 510	CAT His	CCA Pro	1536
GCC Ala	AAG Lys 515	TTG Leu	CGC Arg	GCC Ala	GAG Glu	GCG Ala	CCG Pro 520	CAG Gln	TCG Ser	TCC Ser	GCT Ala	CTG Leu 525	ACG Thr	CGG Arg	ACT Thr	1584
GCC Ala	GTT Val 530	GCC Ala	AGA Arg	GGA Gly	ACA Thr	TCG Ser 535	GGA Gly	TTC Phe	GCA Ala	GAA Glu	TTG Leu 540	CTC Leu	CAC His	GCG Ala	CTG Leu	1632
CAC His 545	CTC Leu	GAT Asp	AGC Ser	TTA Leu	AAT Asn 550	TTA Leu	ATT Ile	CCG Pro	GCG Ala	ATT Ile 555	AAC Asn	TGT Cys	TCA Ser	AAG Lys	ATT Ile 560	1680
ACA Thr	GCC Ala	GAC Asp	AAG Lys 565	ATA Ile	ATA Ile	GCT Ala	ACG Thr	GTA Val	CCC Pro 570	TTG Leu	CCT Pro	CAC His	GTC Val	ACG Thr 575	TAT Tyr	1728
ATC Ile	ATC Ile	AGT Ser 580	TCC Ser	GAA Glu	GCA Ala	CTC Leu	TCG Ser 585	AAC Asn	GCT Ala	GTT Val	GTC Val	TAC Tyr	GAG Glu 590	GTG Val	TCG Ser	1776
GAG Glu	ATC Ile	TTC Phe 595	CTC Leu	AAG Lys	AGT Ser	GCC Ala	ATG Met 600	TTT Phe	ATA Ile	TCT Ser	GCT Ala	ATC Ile 605	AAA Lys	CCC Pro	GAT Asp	1824
TGC Cys 610	TCC Ser	GGC Gly	TTT Phe	AAC Asn	TTT Phe	TCT Ser 615	CAG Gln	ATT Ile	GAT Asp	AGG Arg	CAC His 620	ATT Ile	CCC Pro	ATA Ile	GTC Val	1872
TAC Tyr 625	AAC Asn	ATC Ile	AGC Ser	ACA Thr	CCA Pro 630	AGA Arg	AGA Arg	GGT Gly	TGC Cys	CCC Pro 635	CTT Leu	TGT Cys	GAC Asp	TCT Ser	GTA Val 640	1920
ATC Ile	ATG Met	AGC Ser	TAC Tyr	GAT Asp 645	GAG Glu	AGC Ser	GAT Asp	GGC Gly	CTG Leu 650	CAG Gln	TCT Ser	CTC Leu	ATG Met	TAT Tyr 655	GTC Val	1968
ACT Thr	AAT Asn	GAA Glu 660	AGG Arg	GTG Val	CAG Gln	ACC Thr	AAC Asn 665	CTC Leu	TTT Phe	TTA Leu	GAT Asp	AAG Lys	TCA Ser 670	CCT Pro	TTC Phe	2016

Asn	Arg	Ser	Lys	Glu	Ala	Asn	Glu	Thr	Ala	Ser	His	Leu	Leu	Phe	Gly
210						215					220				
Leu	Pro	Asp	Ser	Leu	Pro	Ser	Leu	Lys	Gly	His	Ala	Thr	Tyr	Asp	Glu
225					230					235					240
Leu	Thr	Phe	Ala	Arg	Asn	Ala	Lys	Tyr	Ala	Leu	Val	Ala	Ile	Leu	Pro
				245					250					255	
Lys	Asp	Ser	Tyr	Gln	Thr	Leu	Leu	Thr	Glu	Asn	Tyr	Thr	Arg	Ile	Phe
			260					265					270		
Leu	Asn	Met	Thr	Glu	Ser	Thr	Pro	Leu	Glu	Phe	Thr	Arg	Thr	Ile	Gln
		275					280					285			
Thr	Arg	Ile	Val	Ser	Ile	Glu	Ala	Arg	Arg	Ala	Cys	Ala	Ala	Gln	Glu
	290					295					300				
Ala	Ala	Pro	Asp	Ile	Phe	Leu	Val	Leu	Phe	Gln	Met	Leu	Val	Ala	His
305					310					315					320
Phe	Leu	Val	Ala	Arg	Gly	Ile	Ala	Glu	His	Arg	Phe	Val	Glu	Val	Asp
				325					330					335	
Cys	Val	Cys	Arg	Gln	Tyr	Ala	Glu	Leu	Tyr	Phe	Leu	Arg	Arg	Ile	Ser
			340					345					350		
Arg	Leu	Cys	Met	Pro	Thr	Phe	Thr	Thr	Val	Gly	Tyr	Asn	His	Thr	Thr
		355					360					365			
Leu	Gly	Ala	Val	Ala	Ala	Thr	Gln	Ile	Ala	Arg	Val	Ser	Ala	Thr	Lys
	370					375					380				
Leu	Ala	Ser	Leu	Pro	Arg	Ser	Ser	Gln	Glu	Thr	Val	Leu	Ala	Met	Val
385					390					395					400
Gln	Leu	Gly	Ala	Arg	Asp	Gly	Ala	Val	Pro	Ser	Ser	Ile	Leu	Glu	Gly
				405					410					415	
Ile	Ala	Met	Val	Val	Glu	His	Met	Tyr	Thr	Ala	Tyr	Thr	Tyr	Val	Tyr
			420					425					430		
Thr	Leu	Gly	Asp	Thr	Glu	Arg	Lys	Leu	Met	Leu	Asp	Ile	His	Thr	Val
		435					440					445			
Leu	Thr	Asp	Ser	Cys	Pro	Pro	Lys	Asp	Ser	Gly	Val	Ser	Glu	Lys	Leu
	450					455					460				
Leu	Arg	Thr	Tyr	Leu	Met	Phe	Thr	Ser	Met	Cys	Thr	Asn	Ile	Glu	Leu
465					470					475					480
Gly	Glu	Met	Ile	Ala	Arg	Phe	Ser	Lys	Pro	Asp	Ser	Leu	Asn	Ile	Tyr
				485					490					495	
Arg	Ala	Phe	Ser	Pro	Cys	Phe	Leu	Gly	Leu	Arg	Tyr	Asp	Leu	His	Pro
			500					505					510		
Ala	Lys	Leu	Arg	Ala	Glu	Ala	Pro	Gln	Ser	Ser	Ala	Leu	Thr	Arg	Thr
		515					520					525			
Ala	Val	Ala	Arg	Gly	Thr	Ser	Gly	Phe	Ala	Glu	Leu	Leu	His	Ala	Leu
	530					535					540				
His	Leu	Asp													

Leu	Phe	Glu	Lys	Phe	Gly	Thr	Pro	Asp	Ser	Ser	Thr	Leu	Pro	Leu	Tyr	
50						55					60					
GCG Ala 65	GCT Ala	AGG Arg	CAC His	CCG Pro	GAA Glu 70	CTA Leu	TCG Ser	TTG Leu	CTA Leu	CGG Arg 75	ATC Ile	ATG Met	CTC Leu	TCA Ser	CCG Pro 80	240
CAC His	CCC Pro	TAC Tyr	GCG Ala	TTA Leu 85	AGA Arg	AGC Ser	CAC His	TTG Leu	TGC Cys 90	GTA Val	GGC Gly	GAA Glu	GAG Glu	ACC Thr 95	GCA Ala	288
TCT Ser	CTT Leu	GGC Gly	GTT Val 100	TAC Tyr	CTG Leu	CAC His	TCC Ser	AAG Lys 105	CCA Pro	GTC Val	GTA Val	CGC Arg	GGC Gly 110	CAC His	GAA Glu	336
TTC Phe	GAG Glu	GAC Asp 115	ACG Thr	CAG Gln	ATA Ile	CTA Leu	CCG Pro 120	GAG Glu	TGC Cys	CGG Arg	CTG Leu	GCC Ala 125	ATA Ile	ACG Thr	AGC Ser	384
GAC Asp	CAG Gln 130	TCT Ser	TAT Tyr	ACC Thr	AAC Asn	TTT Phe 135	AAG Lys	ATT Ile	ATA Ile	GAT Asp	CTG Leu 140	CCA Pro	GCG Ala	GGA Gly	TGC Cys	432
CGT Arg 145	CGC Arg	GTC Val	CCC Pro	ATA Ile	CAC His 150	GCC Ala	GCG Ala	AAC Asn	AAG Lys	CGT Arg 155	GTC Val	GTC Val	ATC Ile	GAC Asp	GAG Glu 160	480
GCC Ala	GCC Ala	AAC Asn	CGC Arg	ATA Ile 165	AAG Lys	GTG Val	TTT Phe	GAC Asp	CCA Pro 170	GAG Glu	TCG Ser	CCT Pro	TTA Leu	CCG Pro 175	CGT Arg	528
CAC His	CCC Pro	ATA Ile	ACA Thr 180	CCC Pro	CGT Arg	GCC Ala	GGT Gly	CAG Gln 185	ACC Thr	AGA Arg	TCT Ser	ATA Ile	CTG Leu 190	AAA Lys	CAC His	576
AAC Asn	ATC Ile	GCA Ala 195	CAG Gln	GTT Val	TGC Cys	GAA Glu	CGG Arg 200	GAT Asp	ATC Ile	GTG Val	TCA Ser	CTT Leu 205	AAC Asn	ACA Thr	GAC Asp	624
AAC Asn	GAG Glu 210	GCC Ala	GCG Ala	TCT Ser	ATG Met	TTC Phe 215	TAC Tyr	ATG Met	ATT Ile	GGA Gly	CTC Leu 220	AGG Arg	CGG Arg	CCG Pro	AGA Arg	672
CTC Leu 225	GGA Gly	GAA Glu	AGC Ser	CCG Pro	GTC Val 230	TGT Cys	GAC Asp	TTC Phe	AAC Asn	ACC Thr 235	GTT Val	ACC Thr	ATC Ile	ATG Met	GAG Glu 240	720
CGT Arg	GCT Ala	AAC Asn	AAC Asn	TCG Ser 245	ATA Ile	ACT Thr	TTT Phe	CTA Leu	CCC Pro 250	AAG Lys	CTA Leu	AAA Lys	CTG Leu	AAC Asn 255	CGG Arg	768
CTA Leu	CAA Gln	CAC His	CTG Leu 260	TTC Phe	CTG Leu	AAG Lys	CAC His	GTG Val 265	TTG Leu	CTG Leu	CGC Arg	AGC Ser	ATG Met 270	GGG Gly	CTG Leu	816
GAA Glu	AAC Asn	ATC Ile 275	GTG Val	TCG Ser	TGT Cys	TTC Phe	TCA Ser 280	TCG Ser	CTG Leu	TAC Tyr	GGC Gly	GCA Ala 285	GAA Glu	CTT Leu	GCC Ala	864
CCT Pro	GCG Ala 290	AAA Lys	ACA Thr	CAC His	GAG Glu	CGG Arg 295	GAG Glu	TTC Phe	TTC Phe	GGC Gly	GCT Ala 300	CTG Leu	CTA Leu	GAA Glu	AGA Arg	912
CTC Leu 305	AAA Lys	CGT Arg	CGG Arg	GTG Val 310	GAG Glu	GAC Asp	GCG Ala	GTC Val	TTC Phe	TGC Cys 315	CTG Leu	AAT Asn	ACC Thr	ATA Ile	GAG Glu 320	960

218

GAT	TTC	CCG	TTT	AGG	GAA	CCC	ATT	CGC	CAA	CCC	CCA	GAT	TGT	TCC	AAG	1008
Asp	Phe	Pro	Phe	Arg	Glu	Pro	Ile	Arg	Gln	Pro	Pro	Asp	Cys	Ser	Lys	
				325					330					335		
GTG	CTT	ATA	GAA	GCC	ATG	GAA	AAG	TAC	TTT	ATG	ATG	TGT	AGC	CCC	AAA	1056
Val	Leu	Ile	Glu	Ala	Met	Glu	Lys	Tyr	Phe	Met	Met	Cys	Ser	Pro	Lys	
			340					345					350			
GAC	CGT	CAA	AGC	GCC	GCA	TGG	CTA	GGT	GCA	GGG	GTG	GTC	GAA	CTG	ATA	1104
Asp	Arg	Gln	Ser	Ala	Ala	Trp	Leu	Gly	Ala	Gly	Val	Val	Glu	Leu	Ile	
		355					360					365				
TGT	GAC	GGC	AAT	CCA	CTT	TCT	GAG	GTG	CTC	GGA	TTT	CTT	GCC	AAG	TAT	1152
Cys	Asp	Gly	Asn	Pro	Leu	Ser	Glu	Val	Leu	Gly	Phe	Leu	Ala	Lys	Tyr	
	370					375					380					
ATG	CCC	ATA	CAA	AAA	GAA	TGC	ACA	GGA	AAC	CTT	TTA	AAA	ATC	TAC	GCT	1200
Met	Pro	Ile	Gln	Lys	Glu	Cys	Thr	Gly	Asn	Leu	Leu	Lys	Ile	Tyr	Ala	
	385				390				395						400	
TTA	TTG	ACC	GTC	TAA												1215
Leu	Leu	Thr	Val													

(2) INFORMATION FOR SEQ ID NO:19:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 404 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

Met	Leu	Arg	Val	Pro	Asp	Val	Lys	Ala	Ser	Leu	Val	Glu	Gly	Ala	Ala	
1				5					10					15		
Arg	Leu	Ser	Thr	Gly	Glu	Arg	Val	Phe	His	Val	Leu	Thr	Ser	Pro	Ala	
			20					25					30			
Val	Ala	Ala	Met	Val	Gly	Val	Ser	Asn	Pro	Glu	Val	Pro	Met	Pro	Leu	
			35				40					45				
Leu	Phe	Glu	Lys	Phe	Gly	Thr	Pro	Asp	Ser	Ser	Thr	Leu	Pro	Leu	Tyr	
	50					55					60					
Ala	Ala	Arg	His	Pro	Glu	Leu	Ser	Leu	Leu	Arg	Ile	Met	Leu	Ser	Pro	
	65				70					75				80		
His	Pro	Tyr	Ala	Leu	Arg	Ser	His	Leu	Cys	Val	Gly	Glu	Glu	Thr	Ala	
				85					90					95		
Ser	Leu	Gly	Val	Tyr	Leu	His	Ser	Lys	Pro	Val	Val	Arg	Gly	His	Glu	
			100					105					110			
Phe	Glu	Asp	Thr	Gln	Ile	Leu	Pro	Glu	Cys	Arg	Leu	Ala	Ile	Thr	Ser	
		115					120					125				
Asp	Gln	Ser	Tyr	Thr	Asn	Phe	Lys	Ile	Ile	Asp	Leu	Pro	Ala	Gly	Cys	
	130					135					140					
Arg	Arg	Val	Pro	Ile	His	Ala	Ala	Asn	Lys	Arg	Val	Val	Ile	Asp	Glu	
	145				150					155				160		
Ala	Ala	Asn	Arg	Ile	Lys	Val	Phe	Asp	Pro	Glu	Ser	Pro	Leu	Pro	Arg	

006290-6240950

220

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

ATG	GCA	GCG	CTC	GAG	GGC	CCC	CTA	CTA	CTG	CCA	CCG	AGC	GCC	TCC	CTG	48
Met	Ala	Ala	Leu	Glu	Gly	Pro	Leu	Leu	Leu	Pro	Pro	Ser	Ala	Ser	Leu	
1				5					10					15		
ACG	ACG	AGT	CCG	CAG	ACC	ACG	TGT	TAT	CAA	GCG	ACT	TGG	GAA	TCA	CAG	96
Thr	Thr	Ser	Pro	Gln	Thr	Thr	Cys	Tyr	Gln	Ala	Thr	Trp	Glu	Ser	Gln	
			20					25					30			
CTG	GAA	ATA	TTC	TGC	TGT	CTG	GCC	ACC	AAC	TCG	CAC	CTG	CAG	GCA	GAG	144
Leu	Glu	Ile	Phe	Cys	Cys	Leu	Ala	Thr	Asn	Ser	His	Leu	Gln	Ala	Glu	
		35					40					45				
CTG	ACC	TTA	GAA	GGT	CTT	GAT	AAG	ATG	ATG	CAG	CCC	GAG	CCC	ACC	TTT	192
Leu	Thr	Leu	Glu	Gly	Leu	Asp	Lys	Met	Met	Gln	Pro	Glu	Pro	Thr	Phe	
	50					55					60					
TTC	GCC	TGC	AGA	GCG	ATA	CGC	AGA	CTA	CTC	CTG	GGG	GAA	CGC	CTC	CAC	240
Phe	Ala	Cys	Arg	Ala	Ile	Arg	Arg	Leu	Leu	Leu	Gly	Glu	Arg	Leu	His	
	65				70					75					80	
CCT	TTT	ATA	CAT	CAA	GAA	GGG	ACT	CTT	TTG	GGA	AAA	GTG	GGT	CGA	CGG	288
Pro	Phe	Ile	His	Gln	Glu	Gly	Thr	Leu	Leu	Gly	Lys	Val	Gly	Arg	Arg	
				85					90					95		
TAC	AGC	GGC	GAA	GGT	TTA	ATA	ATT	GAC	GGT	GGT	GGA	GTG	TTT	ACG	CGC	336
Tyr	Ser	Gly	Glu	Gly	Leu	Ile	Ile	Asp	Gly	Gly	Gly	Val	Phe	Thr	Arg	
			100					105					110			
GGA	CAG	ATA	GAC	ACC	GAC	AAC	TAC	CTA	CCT	GCG	GTG	GGA	TCA	TGG	GAA	384
Gly	Gln	Ile	Asp	Thr	Asp	Asn	Tyr	Leu	Pro	Ala	Val	Gly	Ser	Trp	Glu	
		115					120					125				
CTT	ACC	GAT	GAT	TGT	GAT	AAA	CCC	TGC	GAA	TTC	AGG	GAG	CTA	CGC	TCG	432
Leu	Thr	Asp	Asp	Cys	Asp	Lys	Pro	Cys	Glu	Phe	Arg	Glu	Leu	Arg	Ser	
	130					135					140					
CTG	TAT	CTT	CCC	GCG	CTA	CTA	ACG	TGC	ACC	ATA	TGT	TAC	AAA	GCC	ATG	480
Leu	Tyr	Leu	Pro	Ala	Leu	Leu	Thr	Cys	Thr	Ile	Cys	Tyr	Lys	Ala	Met	
	145				150					155					160	
TTC	AGG	ATA	GTG	TGC	AGG	TAC	CTG	GAG	TTC	TGG	GAG	TTC	GAA	CAG	TGT	528
Phe	Arg	Ile	Val	Cys	Arg	Tyr	Leu	Glu	Phe	Trp	Glu	Phe	Glu	Gln	Cys	
				165					170					175		
TTT	CAT	GCG	TTT	CTG	GCG	GTG	TTG	CCC	CAT	AGT	CTA	CAA	CCC	ACA	ATC	576
Phe	His	Ala	Phe	Leu	Ala	Val	Leu	Pro	His	Ser	Leu	Gln	Pro	Thr	Ile	
			180					185					190			
TAT	CAA	AAT	TAT	TTT	GCA	CTC	CTG	GAG	AGC	CTG	AAG	CAT	CTC	TCG	TTT	624
Tyr	Gln	Asn	Tyr	Phe	Ala	Leu	Leu	Glu	Ser	Leu	Lys	His	Leu	Ser	Phe	
		195					200					205				
TCA	ATA	ATG	CCA	CCC	GCA	TCC	CCA	GAC	GCA	CAG	CTA	CAT	TTT	TTA	AAG	672
Ser	Ile	Met	Pro	Pro	Ala	Ser	Pro	Asp	Ala	Gln	Leu	His	Phe	Leu	Lys	
	210					215					220					
TTT	AAC	ATC	AGC	AGC	TTC	ATG	GCC	ACG	TGG	GGG	TGG	CAC	GGA	GAG	CTG	720
Phe	Asn	Ile	Ser	Ser	Phe	Met	Ala	Thr	Trp	Gly	Trp	His	Gly	Glu	Leu	
	225				230					235					240	
GTC	TCG	CTG	CGC	CGT	GCC	ATC	GCT	CAC	AAC	GTA	GAG	CGA	CTG	CCC	ACC	768
Val	Ser	Leu	Arg	Arg	Ala	Ile	Ala	His	Asn	Val	Glu	Arg	Leu	Pro	Thr	
				245					250					255		
GTG	CTG	AAG	AAC	CTG	TCG	AAA	CAG	AGT	AAG	CAC	CAG	GAC	GTC	AAG	GTT	816

00697179-000000

Val	Leu	Lys	Asn	Leu	Ser	Lys	Gln	Ser	Lys	His	Gln	Asp	Val	Lys	Val	
			260					265					270			
AAC	GGA	CGG	GAT	CTG	GTG	GGC	TTT	CAG	CTG	GCT	CTA	AAC	CAG	CTC	GTG	864
Asn	Gly	Arg	Asp	Leu	Val	Gly	Phe	Gln	Leu	Ala	Leu	Asn	Gln	Leu	Val	
		275					280					285				
TCC	CGT	CTG	CAC	GTA	AAA	ATC	CAA	CGC	AAG	GAC	CCC	GGA	CCA	AAG	CCA	912
Ser	Arg	Leu	His	Val	Lys	Ile	Gln	Arg	Lys	Asp	Pro	Gly	Pro	Lys	Pro	
		290				295					300					
TAC	AGG	GTG	GTC	GTC	AGT	ACC	CCA	GAT	TGT	ACC	TAC	TAT	CTA	GTG	TAT	960
Tyr	Arg	Val	Val	Val	Ser	Thr	Pro	Asp	Cys	Thr	Tyr	Tyr	Leu	Val	Tyr	
		305			310					315					320	
CCG	GGC	ACA	CCG	GCC	ATC	TAC	AGA	CTC	GTC	ATG	TGT	ATG	GCA	GTG	GCA	1008
Pro	Gly	Thr	Pro	Ala	Ile	Tyr	Arg	Leu	Val	Met	Cys	Met	Ala	Val	Ala	
				325					330					335		
GAC	TGC	ATC	GGC	CAC	TCG	TGC	AGC	GGA	CTG	CAC	CCC	TGC	GCA	AAC	TTT	1056
Asp	Cys	Ile	Gly	His	Ser	Cys	Ser	Gly	Leu	His	Pro	Cys	Ala	Asn	Phe	
			340					345					350			
TTA	GGC	ACC	CAC	GAG	ACA	CCG	CGT	CTC	CTG	GCG	GCG	ACG	CTT	TCA	AGA	1104
Leu	Gly	Thr	His	Glu	Thr	Pro	Arg	Leu	Leu	Ala	Ala	Thr	Leu	Ser	Arg	
		355					360					365				
ATC	CGG	TAC	GCG	CCG	AAA	GAC	CGG	CGA	GCA	GCC	ATG	AAA	GGA	AAT	TTG	1152
Ile	Arg	Tyr	Ala	Pro	Lys	Asp	Arg	Arg	Ala	Ala	Met	Lys	Gly	Asn	Leu	
		370				375					380					
CAG	GCG	TGC	TTC	CAA	CGA	TAC	GCG	GCC	ACG	GAC	GCG	CGG	ACT	CTG	GGC	1200
Gln	Ala	Cys	Phe	Gln	Arg	Tyr	Ala	Ala	Thr	Asp	Ala	Arg	Thr	Leu	Gly	
					390					395					400	
AGC	TCT	ACA	GTG	TCA	GAC	ATG	CTG	GAA	CCC	ACA	AAA	CAC	GTC	AGT	TTG	1248
Ser	Ser	Thr	Val	Ser	Asp	Met	Leu	Glu	Pro	Thr	Lys	His	Val	Ser	Leu	
				405					410					415		
GAA	AAC	TTC	AAG	ATC	ACC	ATA	TTC	AAC	ACC	AAC	ATG	GTG	ATT	AAC	ACT	1296
Glu	Asn	Phe	Lys	Ile	Thr	Ile	Phe	Asn	Thr	Asn	Met	Val	Ile	Asn	Thr	
			420					425					430			
AAG	ATA	AGC	TGC	CAC	GTT	CCT	AAC	ACC	CTG	CAA	AAG	ACT	ATT	TTA	AAC	1344
Lys	Ile	Ser	Cys	His	Val	Pro	Asn	Thr	Leu	Gln	Lys	Thr	Ile	Leu	Asn	
		435					440					445				
ATC	CCC	AGA	TTG	ACC	AAC	AAT	TTT	GTT	ATA	CGA	AAG	TAC	TCC	GTA	AAG	1392
Ile	Pro	Arg	Leu	Thr	Asn	Asn	Phe	Val	Ile	Arg	Lys	Tyr	Ser	Val	Lys	
		450				455					460					
GAA	CCT	TCT	TTT	ACC	ATA	AGC	GTG	TTT	TTT	TCC	GAC	AAC	ATG	TGT	CAA	1440
Glu	Pro	Ser	Phe	Thr	Ile	Ser	Val	Phe	Phe	Ser	Asp	Asn	Met	Cys	Gln	
					470					475					480	
GGC	ACC	GCA	ATA	AAC	ATC	AAC	ATC	AGT	GGG	GAC	ATG	CTG	CAC	TTT	CTC	1488
Gly	Thr	Ala	Ile	Asn	Ile	Asn	Ile	Ser	Gly	Asp	Met	Leu	His	Phe	Leu	
				485					490					495		
TTC	GCA	ATG	GGT	ACG	CTG	AAA	TGC	TTT	CTG	CCA	ATC	AGG	CAC	ATA	TTT	1536
Phe	Ala	Met	Gly	Thr	Leu	Lys	Cys	Phe	Leu	Pro	Ile	Arg	His	Ile	Phe	
			500					505					510			
CCT	GTA	TCG	ATA	GCA	AAT	TGG	AAC	TCC	ACG	TTG	GAC	CTG	CAC	GGA	CTG	1584
Pro	Val	Ser	Ile	Ala	Asn	Trp	Asn	Ser	Thr	Leu	Asp	Leu	His	Gly	Leu	
			515				520					525				

222

GAA Glu 530	AAC Asn	CAG Gln	TAC Tyr	ATG Met	GTG Val	AGA Arg	ATG Met	GGG Gly	CGA Arg	AAA Lys	AAC Asn	GTA Val	TTT Phe	TGG Trp	ACC Thr	1632
ACA Thr 545	AAC Asn	TTT Phe	CCA Pro	TCT Ser	GTG Val	GTC Val	TCC Ser	AGC Ser	AAG Lys	GAT Asp	GGG Gly	CTA Leu	AAC Asn	GTG Val	TCC Ser	1680
TGG Trp	TTT Phe	AAG Lys	GCC Ala	GCG Ala	ACA Thr	GCC Ala	ACG Thr	ATT Ile	TCT Ser	AAA Lys	GTG Val	TAC Tyr	GGG Gly	CAG Gln	CCT Pro	1728
CTT Leu	GTG Val	GAA Glu	CAG Gln	ATT Ile	CGC Arg	CAC His	GAG Glu	CTG Leu	GCG Ala	CCC Pro	ATT Ile	CTC Leu	ACG Thr	GAC Asp	CAG Gln	1776
CAC His	GCG Ala	CGC Arg	ATC Ile	GAC Asp	GGA Gly	AAC Asn	AAA Lys	AAT Asn	AGA Arg	ATA Ile	TTC Phe	TCC Ser	CTA Leu	CTT Leu	GAG Glu	1824
CAC His	AGA Arg	AAC Asn	CGT Arg	TCC Ser	CAA Gln	ATA Ile	CAG Gln	ACG Thr	CTA Leu	CAC His	AAA Lys	AGG Arg	TTC Phe	CTG Leu	GAG Glu	1872
TGT Cys	CTG Leu	GTG Val	GAA Glu	TGC Cys	TGT Cys	TCG Ser	TTT Phe	CTC Leu	AGG Arg	CTT Leu	GAC Asp	GTG Val	GCT Ala	TGC Cys	ATT Ile	1920
AGG Arg	CGA Arg	GCC Ala	GCC Ala	GCC Ala	CGG Arg	GGC Gly	CTG Leu	TTT Phe	GAC Asp	TTC Phe	TCA Ser	AAG Lys	AAG Lys	ATA Ile	ATC Ile	1968
AGT Ser	CAC His	ACT Thr	AAA Lys	AGC Ser	AAA Lys	CAC His	GAG Glu	TGC Cys	GCA Ala	GTA Val	CTG Leu	GGA Gly	TAT Tyr	AAA Lys	AAG Lys	2016
TGT Cys	AAC Asn	CTA Leu	ATC Ile	CCG Pro	AAA Lys	ATC Ile	TAT Tyr	GCC Ala	CGA Arg	AAC Asn	AAG Lys	AAG Lys	ACC Thr	AGG Arg	CTA Leu	2064
GAC Asp	GAG Glu	TTG Leu	GGC Gly	CGC Arg	AAT Asn	GCA Ala	AAC Asn	TTC Phe	ATT Ile	TCG Ser	TTC Phe	GTC Val	GCC Ala	ACC Thr	ACG Thr	2112
GGT Gly	CAT His	CGG Arg	TTC Phe	GCC Ala	GCT Ala	CTA Leu	AAG Lys	CCA Pro	CAA Gln	ATT Ile	GTC Val	CGT Arg	CAC His	GCC Ala	ATT Ile	2160
CGC Arg	AAA Lys	CTA Leu	GGC Gly	CTG Leu	CAC His	TGG Trp	CGC Arg	CAC His	CGA Arg	ACG Thr	GCC Ala	GCG Ala	TCC Ser	AAC Asn	GAG Glu	2208
CAG Gln	ACA Thr	CCG Pro	CCA Pro	GCC Ala	GAT Asp	CCC Pro	CGC Arg	GTA Val	CGT Arg	TGC Cys	GTC Val	CGT Arg	CCG Pro	CTG Leu	GTC Val	2256
TAA																2259

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 752 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

223

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

```

Met Ala Ala Leu Glu Gly Pro Leu Leu Leu Pro Pro Ser Ala Ser Leu
 1           5           10
Thr Thr Ser Pro Gln Thr Thr Cys Tyr Gln Ala Thr Trp Glu Ser Gln
          20           25           30
Leu Glu Ile Phe Cys Cys Leu Ala Thr Asn Ser His Leu Gln Ala Glu
          35           40           45
Leu Thr Leu Glu Gly Leu Asp Lys Met Met Gln Pro Glu Pro Thr Phe
          50           55           60
Phe Ala Cys Arg Ala Ile Arg Arg Leu Leu Leu Gly Glu Arg Leu His
 65           70           75
Pro Phe Ile His Gln Glu Gly Thr Leu Leu Gly Lys Val Gly Arg Arg
          85           90           95
Tyr Ser Gly Glu Gly Leu Ile Ile Asp Gly Gly Gly Val Phe Thr Arg
          100          105          110
Gly Gln Ile Asp Thr Asp Asn Tyr Leu Pro Ala Val Gly Ser Trp Glu
          115          120          125
Leu Thr Asp Asp Cys Asp Lys Pro Cys Glu Phe Arg Glu Leu Arg Ser
          130          135          140
Leu Tyr Leu Pro Ala Leu Leu Thr Cys Thr Ile Cys Tyr Lys Ala Met
          145          150          155          160
Phe Arg Ile Val Cys Arg Tyr Leu Glu Phe Trp Glu Phe Glu Gln Cys
          165          170          175
Phe His Ala Phe Leu Ala Val Leu Pro His Ser Leu Gln Pro Thr Ile
          180          185          190
Tyr Gln Asn Tyr Phe Ala Leu Leu Glu Ser Leu Lys His Leu Ser Phe
          195          200          205
Ser Ile Met Pro Pro Ala Ser Pro Asp Ala Gln Leu His Phe Leu Lys
          210          215          220
Phe Asn Ile Ser Ser Phe Met Ala Thr Trp Gly Trp His Gly Glu Leu
          225          230          235          240
Val Ser Leu Arg Arg Ala Ile Ala His Asn Val Glu Arg Leu Pro Thr
          245          250          255
Val Leu Lys Asn Leu Ser Lys Gln Ser Lys His Gln Asp Val Lys Val
          260          265          270
Asn Gly Arg Asp Leu Val Gly Phe Gln Leu Ala Leu Asn Gln Leu Val
          275          280          285
Ser Arg Leu His Val Lys Ile Gln Arg Lys Asp Pro Gly Pro Lys Pro
          290          295          300
Tyr Arg Val Val Val Ser Thr Pro Asp Cys Thr Tyr Tyr Leu Val Tyr
          305          310          315          320
Pro Gly Thr Pro Ala Ile Tyr Arg Leu Val Met Cys Met Ala Val Ala
          325          330          335

```

006690.6272066

690

695

700

Gly	His	Arg	Phe	Ala	Ala	Leu	Lys	Pro	Gln	Ile	Val	Arg	His	Ala	Ile
705					710					715					720
Arg	Lys	Leu	Gly	Leu	His	Trp	Arg	His	Arg	Thr	Ala	Ala	Ser	Asn	Glu
				725					730					735	
Gln	Thr	Pro	Pro	Ala	Asp	Pro	Arg	Val	Arg	Cys	Val	Arg	Pro	Leu	Val
			740					745					750		

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 364 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(11) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(ix) FEATURE:

- (A) NAME/KEY: CDS
(B) LOCATION: 1..364
(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

[illegible]

226

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 121 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

Met Val Arg Pro Thr Glu Ala Glu Val Lys Lys Ser Leu Ser Arg Leu
 1 5 10 15
 Pro Ala Ala Arg Lys Arg Ala Gly Asn Arg Ala His Leu Ala Thr Tyr
 20 25 30
 Arg Arg Leu Leu Lys Tyr Ser Thr Leu Pro Asp Leu Trp Arg Phe Leu
 35 40 45
 Ser Ser Arg Pro Gln Asn Pro Pro Leu Gly His His Arg Leu Phe Phe
 50 55 60
 Glu Val Thr Leu Gly His Arg Ile Ala Asp Cys Val Ile Leu Val Ser
 65 70 75 80
 Gly Gly His Gln Pro Val Cys Tyr Val Val Glu Leu Lys Thr Cys Leu
 85 90 95
 Ser His Gln Leu Ile Pro Thr Asn Thr Val Arg Thr Ser Gln Arg Ala
 100 105 110
 Gln Gly Leu Cys Gln Leu Ser Asp Ser
 115 120

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 918 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..918
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

ATG GCA CTC GAC AAG AGT ATA GTG GTT AAC TTC ACC TCC AGA CTC TTC
 Met Ala Leu Asp Lys Ser Ile Val Val Asn Phe Thr Ser Arg Leu Phe
 1 5 10 15 48
 GCT GAT GAA CTG GCC GCC CTT CAG TCA AAA ATA GGG AGC GTA CTG CCG
 Ala Asp Glu Leu Ala Ala Leu Gln Ser Lys Ile Gly Ser Val Leu Pro
 20 25 30 96

227

CTC	GGA	GAT	TGC	CAC	CGT	TTA	CAA	AAT	ATA	CAG	GCA	TTG	GGC	CTG	GGG	144
Leu	Gly	Asp	Cys	His	Arg	Leu	Gln	Asn	Ile	Gln	Ala	Leu	Gly	Leu	Gly	
		35					40					45				
TGC	GTA	TGC	TCA	CGT	GAG	ACA	TCT	CCG	GAC	TAC	ATC	CAA	ATT	ATG	CAG	192
Cys	Val	Cys	Ser	Arg	Glu	Thr	Ser	Pro	Asp	Tyr	Ile	Gln	Ile	Met	Gln	
	50				55					60						
TAT	CTA	TCC	AAG	TGC	ACA	CTC	GCT	GTC	CTG	GAG	GAG	GTT	CGC	CCG	GAC	240
Tyr	Leu	Ser	Lys	Cys	Thr	Leu	Ala	Val	Leu	Glu	Glu	Val	Arg	Pro	Asp	
65					70					75					80	
AGC	CTG	CGC	CTA	ACG	CGG	ATG	GAT	CCC	TCT	GAC	AAC	CTT	CAG	ATA	AAA	288
Ser	Leu	Arg	Leu	Thr	Arg	Met	Asp	Pro	Ser	Asp	Asn	Leu	Gln	Ile	Lys	
				85				90						95		
AAC	GTA	TAT	GCC	CCC	TTT	TTT	CAG	TGG	GAC	AGC	AAC	ACC	CAG	CTA	GCA	336
Asn	Val	Tyr	Ala	Pro	Phe	Phe	Gln	Trp	Asp	Ser	Asn	Thr	Gln	Leu	Ala	
			100					105					110			
GTG	CTA	CCC	CCA	TTT	TTT	AGC	CGA	AAG	GAT	TCC	ACC	ATT	GTG	CTC	GAA	384
Val	Leu	Pro	Pro	Phe	Phe	Ser	Arg	Lys	Asp	Ser	Thr	Ile	Val	Leu	Glu	
		115					120					125				
TCC	AAC	GGA	TTT	GAC	CCC	GTG	TTC	CCC	ATG	GTC	GTG	CCG	CAG	CAA	CTG	432
Ser	Asn	Gly	Phe	Asp	Pro	Val	Phe	Pro	Met	Val	Val	Pro	Gln	Gln	Leu	
	130					135					140					
GGG	CAC	GCT	ATT	CTG	CAG	CAG	CTG	TTG	GTG	TAC	CAC	ATC	TAC	TCC	AAA	480
Gly	His	Ala	Ile	Leu	Gln	Gln	Leu	Leu	Val	Tyr	His	Ile	Tyr	Ser	Lys	
145					150					155					160	
ATA	TCG	GCC	GGG	GCC	CCG	GAT	GAT	GTA	AAT	ATG	GCG	GAA	CTT	GAT	CTA	528
Ile	Ser	Ala	Gly	Ala	Pro	Asp	Asp	Val	Asn	Met	Ala	Glu	Leu	Asp	Leu	
				165					170					175		
TAT	ACC	ACC	AAT	GTG	TCA	TTT	ATG	GGG	CGC	ACA	TAT	CGT	CTG	GAC	GTA	576
Tyr	Thr	Thr	Asn	Val	Ser	Phe	Met	Gly	Arg	Thr	Tyr	Arg	Leu	Asp	Val	
			180					185					190			
GAC	AAC	ACG	GAT	CCA	CGT	ACT	GCC	CTG	CGA	GTG	CTT	GAC	GAT	CTG	TCC	624
Asp	Asn	Thr	Asp	Pro	Arg	Thr	Ala	Leu	Arg	Val	Leu	Asp	Asp	Leu	Ser	
		195					200					205				
ATG	TAC	CTT	TGT	ATC	CTA	TCA	GCC	TTG	GTT	CCC	AGG	GGG	TGT	CTC	CGT	672
Met	Tyr	Leu	Cys	Ile	Leu	Ser	Ala	Leu	Val	Pro	Arg	Gly	Cys	Leu	Arg	
	210					215					220					
CTG	CTC	ACG	GCG	CTC	GTG	CGG	CAC	GAC	AGG	CAT	CCT	CTG	ACA	GAG	GTG	720
Leu	Leu	Thr	Ala	Leu	Val	Arg	His	Asp	Arg	His	Pro	Leu	Thr	Glu	Val	
225					230					235					240	
TTT	GAG	GGG	GTG	GTG	CCA	GAT	GAG	GTG	ACC	AGG	ATA	GAT	CTC	GAC	CAG	768
Phe	Glu	Gly	Val	Val	Pro	Asp	Glu	Val	Thr	Arg	Ile	Asp	Leu	Asp	Gln	
				245					250					255		
TTG	AGC	GTC	CCA	GAT	GAC	ATC	ACC	AGG	ATG	CGC	GTC	ATG	TTC	TCC	TAT	816
Leu	Ser	Val	Pro	Asp	Asp	Ile	Thr	Arg	Met	Arg	Val	Met	Phe	Ser	Tyr	
			260					265					270			
CTT	CAG	AGT	CTC	AGT	TCT	ATA	TTT	AAT	CTT	GGC	CCC	AGA	CTG	CAC	GTG	864
Leu	Gln	Ser	Leu	Ser	Ser	Ile	Phe	Asn	Leu	Gly	Pro	Arg	Leu	His	Val	
		275					280					285				
TAT	GCC	TAC	TCG	GCA	GAG	ACT	TTG	GCG	GCC	TCC	TGT	TGG	TAT	TCC	CCA	912
Tyr	Ala	Tyr	Ser	Ala	Glu	Thr	Leu	Ala	Ala	Ser	Cys	Trp	Tyr	Ser	Pro	
	290					295					300					

006290-6440960

(2) INFORMATION FOR SEQ ID NO:25:

(A) LENGTH: 305 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(x¹) SEQUENCE DESCRIPTION: SEQ ID NO:25:

Met 1	Ala	Leu	Asp	Lys 5	Ser	Ile	Val	Val	Asn 10	Phe	Thr	Ser	Arg	Leu 15	Phe
Ala	Asp	Glu	Leu 20	Ala	Ala	Leu	Gln	Ser 25	Lys	Ile	Gly	Ser	Val 30	Leu	Pro
Leu	Gly 35	Asp	Cys	His	Arg	Leu	Gln 40	Asn	Ile	Gln	Ala	Leu 45	Gly	Leu	Gly
Cys 50	Val	Cys	Ser	Arg	Glu	Thr 55	Ser	Pro	Asp	Tyr	Ile 60	Gln	Ile	Met	Gln
Tyr 65	Leu	Ser	Lys	Cys	Thr 70	Leu	Ala	Val	Leu	Glu 75	Glu	Val	Arg	Pro	Asp 80
Ser	Leu	Arg	Leu	Thr 85	Arg	Met	Asp	Pro	Ser 90	Asp	Asn	Leu	Gln	Ile 95	Lys
Asn	Val	Tyr	Ala 100	Pro	Phe	Phe	Gln	Trp 105	Asp	Ser	Asn	Thr	Gln 110	Leu	Ala
Val	Leu 115	Pro	Pro	Phe	Phe	Ser	Arg 120	Lys	Asp	Ser	Thr	Ile 125	Val	Leu	Glu
Ser	Asn 130	Gly	Phe	Asp	Pro	Val 135	Phe	Pro	Met	Val	Val 140	Pro	Gln	Gln	Leu
Gly 145	His	Ala	Ile	Leu	Gln 150	Gln	Leu	Leu	Val	Tyr 155	His	Ile	Tyr	Ser	Lys 160
Ile	Ser	Ala	Gly	Ala 165	Pro	Asp	Asp	Val	Asn 170	Met	Ala	Glu	Leu	Asp 175	Leu
Tyr	Thr	Thr	Asn 180	Val	Ser	Phe	Met	Gly 185	Arg	Thr	Tyr	Arg	Leu 190	Asp	Val
Asp	Asn 195	Thr	Asp	Pro	Arg	Thr 200	Ala	Leu	Arg	Val	Leu 205	Asp	Asp	Leu	Ser
Met 210	Tyr	Leu	Cys	Ile	Leu	Ser 215	Ala	Leu	Val	Pro	Arg 220	Gly	Cys	Leu	Arg
Leu 225	Leu	Thr	Ala	Leu	Val 230	Arg	His	Asp	Arg	His 235	Pro	Leu	Thr	Glu	Val 240
Phe	Glu	Gly	Val 245	Val	Pro	Asp	Glu	Val	Thr 250	Arg	Ile	Asp	Leu	Asp 255	Gln
Leu	Ser	Val 260	Pro	Asp	Asp	Ile	Thr	Arg 265	Met	Arg	Val	Met	Phe 270	Ser	Tyr

229

Leu Gln Ser Leu Ser Ser Ile Phe Asn Leu Gly Pro Arg Leu His Val
 275 280 285

Tyr Ala Tyr Ser Ala Glu Thr Leu Ala Ala Ser Cys Trp Tyr Ser Pro
 290 295 300

Arg
 305

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 873 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..873
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

ATG GCG TCA TCT GAT ATT CTG TCG GTT GCA AGG ACG GAT GAC GGC TCC	48
Met Ala Ser Ser Asp Ile Leu Ser Val Ala Arg Thr Asp Asp Gly Ser	
1 5 10 15	
GTC TGT GAA GTC TCC CTG CGT GGA GGT AGG AAA AAA ACT ACC GTC TAC	96
Val Cys Glu Val Ser Leu Arg Gly Gly Arg Lys Lys Thr Thr Val Tyr	
20 25 30	
CTG CCG GAC ACT GAA CCC TGG GTG GTA GAG ACC GAC GCC ATC AAA GAC	144
Leu Pro Asp Thr Glu Pro Trp Val Val Glu Thr Asp Ala Ile Lys Asp	
35 40 45	
GCC TTC CTC AGC GAC GGG ATC GTG GAT ATG GCT CGA AAG CTT CAT CGT	192
Ala Phe Leu Ser Asp Gly Ile Val Asp Met Ala Arg Lys Leu His Arg	
50 55 60	
GGT GCC CTG CCC TCA AAT TCT CAC AAC GGC TTG AGG ATG GTG CTT TTT	240
Gly Ala Leu Pro Ser Asn Ser His Asn Gly Leu Arg Met Val Leu Phe	
65 70 75 80	
TGT TAT TGT TAC TTG CAA AAT TGT GTG TAC CTA GCC CTG TTT CTG TGC	288
Cys Tyr Cys Tyr Leu Gln Asn Cys Val Tyr Leu Ala Leu Phe Leu Cys	
85 90 95	
CCC CTT AAT CCT TAC TTG GTA ACT CCC TCA AGC ATT GAG TTT GCC GAG	336
Pro Leu Asn Pro Tyr Leu Val Thr Pro Ser Ser Ile Glu Phe Ala Glu	
100 105 110	
CCC GTT GTG GCA CCT GAG GTG CTC TTC CCA CAC CCG GCT GAG ATG TCT	384
Pro Val Val Ala Pro Glu Val Leu Phe Pro His Pro Ala Glu Met Ser	
115 120 125	
CGC GGT TGC GAT GAC GCG ATT TTC TGT AAA CTG CCC TAT ACC GTG CCT	432
Arg Gly Cys Asp Asp Ala Ile Phe Cys Lys Leu Pro Tyr Thr Val Pro	
130 135 140	

006230-0210960

230

ATA ATC AAC ACC ACG TTT GGA CGC ATT TAC CCG AAC TCT ACA CGC GAG	480
Ile Ile Asn Thr Thr Phe Gly Arg Ile Tyr Pro Asn Ser Thr Arg Glu	
145 150 155 160	
CCG GAC GGC AGG CCT ACG GAT TAC TCC ATG GCC CTT AGA AGG GCT TTT	528
Pro Asp Gly Arg Pro Thr Asp Tyr Ser Met Ala Leu Arg Arg Ala Phe	
165 170 175	
GCA GTT ATG GTT AAC ACG TCA TGT GCA GGA GTG ACA TTG TGC CGC GGA	576
Ala Val Met Val Asn Thr Ser Cys Ala Gly Val Thr Leu Cys Arg Gly	
180 185 190	
GAA ACT CAG ACC GCA TCC CGT AAC CAC ACT GAG TGG GAA AAT CTG CTG	624
Glu Thr Gln Thr Ala Ser Arg Asn His Thr Glu Trp Glu Asn Leu Leu	
195 200 205	
GCT ATG TTT TCT GTG ATT ATC TAT GCC TTA GAT CAC AAC TGT CAC CCG	672
Ala Met Phe Ser Val Ile Ile Tyr Ala Leu Asp His Asn Cys His Pro	
210 215 220	
GAA GCA CTG TCT ATC GCG AGC GGC ATC TTT GAC GAG CGT GAC TAT GGA	720
Glu Ala Leu Ser Ile Ala Ser Gly Ile Phe Asp Glu Arg Asp Tyr Gly	
225 230 235 240	
TTA TTC ATC TCT CAG CCC CGG AGC GTG CCC TCG CCT ACC CCT TGC GAC	768
Leu Phe Ile Ser Gln Pro Arg Ser Val Pro Ser Pro Thr Pro Cys Asp	
245 250 255	
GTG TCG TGG GAA GAT ATC TAC AAC GGG ACT TAC CTA GCT CGG CCT GGA	816
Val Ser Trp Glu Asp Ile Tyr Asn Gly Thr Tyr Leu Ala Arg Pro Gly	
260 265 270	
AAC TGT GAC CCC TGG CCC AAT CTA TCC ACC CCT CCC TTG ATT CTA AAT	864
Asn Cys Asp Pro Trp Pro Asn Leu Ser Thr Pro Pro Leu Ile Leu Asn	
275 280 285	
TTT AAA TAA	873
Phe Lys	
290	

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 290 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

Met	Ala	Ser	Ser	Asp	Ile	Leu	Ser	Val	Ala	Arg	Thr	Asp	Asp	Gly	Ser
1				5					10					15	
Val	Cys	Glu	Val	Ser	Leu	Arg	Gly	Gly	Arg	Lys	Lys	Thr	Thr	Val	Tyr
			20					25					30		
Leu	Pro	Asp	Thr	Glu	Pro	Trp	Val	Val	Glu	Thr	Asp	Ala	Ile	Lys	Asp
			35				40					45			
Ala	Phe	Leu	Ser	Asp	Gly	Ile	Val	Asp	Met	Ala	Arg	Lys	Leu	His	Arg
	50					55				60					
Gly	Ala	Leu	Pro	Ser	Asn	Ser	His	Asn	Gly	Leu	Arg	Met	Val	Leu	Phe
65					70				75						80

231

Cys Tyr Cys Tyr Leu Gln Asn Cys Val Tyr Leu Ala Leu Phe Leu Cys
 85 90 95
 Pro Leu Asn Pro Tyr Leu Val Thr Pro Ser Ser Ile Glu Phe Ala Glu
 100 105 110
 Pro Val Val Ala Pro Glu Val Leu Phe Pro His Pro Ala Glu Met Ser
 115 120 125
 Arg Gly Cys Asp Asp Ala Ile Phe Cys Lys Leu Pro Tyr Thr Val Pro
 130 135 140
 Ile Ile Asn Thr Thr Phe Gly Arg Ile Tyr Pro Asn Ser Thr Arg Glu
 145 150 155 160
 Pro Asp Gly Arg Pro Thr Asp Tyr Ser Met Ala Leu Arg Arg Ala Phe
 165 170 175
 Ala Val Met Val Asn Thr Ser Cys Ala Gly Val Thr Leu Cys Arg Gly
 180 185 190
 Glu Thr Gln Thr Ala Ser Arg Asn His Thr Glu Trp Glu Asn Leu Leu
 195 200 205
 Ala Met Phe Ser Val Ile Ile Tyr Ala Leu Asp His Asn Cys His Pro
 210 215 220
 Glu Ala Leu Ser Ile Ala Ser Gly Ile Phe Asp Glu Arg Asp Tyr Gly
 225 230 235 240
 Leu Phe Ile Ser Gln Pro Arg Ser Val Pro Ser Pro Thr Pro Cys Asp
 245 250 255
 Val Ser Trp Glu Asp Ile Tyr Asn Gly Thr Tyr Leu Ala Arg Pro Gly
 260 265 270
 Asn Cys Asp Pro Trp Pro Asn Leu Ser Thr Pro Pro Leu Ile Leu Asn
 275 280 285
 Phe Lys
 290

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 363 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..363
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

ATG AGC ATG ACT TTC CCC GTC TCC AGT CAC CGG AGG AAT GGT GGA CGG
 Met Ser Met Thr Phe Pro Val Ser Ser His Arg Arg Asn Gly Gly Arg
 1 5 10 15

232

CTC	CGT	CCT	GGT	GCG	AAT	GGC	CAC	CAA	GCC	TCC	CGT	GAT	TGG	TCT	TAT	96
Leu	Arg	Pro	Gly	Ala	Asn	Gly	His	Gln	Ala	Ser	Arg	Asp	Trp	Ser	Tyr	
			20					25					30			
AAC	AGT	GCT	CTT	CCT	CCT	AGT	CAT	AGG	CGC	CTG	CGT	CTA	CTG	CTG	CAT	144
Asn	Ser	Ala	Leu	Pro	Pro	Ser	His	Arg	Arg	Leu	Arg	Leu	Leu	Leu	His	
		35					40					45				
TCG	CGT	GTT	CCT	GGC	GGC	TCG	ACT	GTG	GCG	CGC	CAC	CCC	ACT	AGG	CAG	192
Ser	Arg	Val	Pro	Gly	Gly	Ser	Thr	Val	Ala	Arg	His	Pro	Thr	Arg	Gln	
	50					55					60					
GGC	CAC	CGT	GGC	GTA	TCA	GGT	CCT	TCG	CAC	CCT	GGG	ACC	GCA	GGC	CGG	240
Gly	His	Arg	Gly	Val	Ser	Gly	Pro	Ser	His	Pro	Gly	Thr	Ala	Gly	Arg	
65				70						75					80	
GTC	ACA	TGC	ACC	GCC	GAC	GGT	GGG	CAT	AGC	TAC	CCA	GGA	GCC	CTA	CCG	288
Val	Thr	Cys	Thr	Ala	Asp	Gly	Gly	His	Ser	Tyr	Pro	Gly	Ala	Leu	Pro	
				85					90					95		
TAC	AAT	ATA	CAT	GCC	AGA	TTA	GAA	CGG	GGT	GTG	TGC	TAT	AAT	GGA	TGG	336
Tyr	Asn	Ile	His	Ala	Arg	Leu	Glu	Arg	Gly	Val	Cys	Tyr	Asn	Gly	Trp	
			100					105					110			
CTA	TGG	GGG	GGG	GCT	GTA	GAT	AAT	TGA								363
Leu	Trp	Gly	Gly	Ala	Val	Asp	Asn									
		115				120										

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 120 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

Met	Ser	Met	Thr	Phe	Pro	Val	Ser	Ser	His	Arg	Arg	Asn	Gly	Gly	Arg	
				5					10					15		
Leu	Arg	Pro	Gly	Ala	Asn	Gly	His	Gln	Ala	Ser	Arg	Asp	Trp	Ser	Tyr	
			20					25					30			
Asn	Ser	Ala	Leu	Pro	Pro	Ser	His	Arg	Arg	Leu	Arg	Leu	Leu	Leu	His	
		35					40					45				
Ser	Arg	Val	Pro	Gly	Gly	Ser	Thr	Val	Ala	Arg	His	Pro	Thr	Arg	Gln	
	50					55					60					
Gly	His	Arg	Gly	Val	Ser	Gly	Pro	Ser	His	Pro	Gly	Thr	Ala	Gly	Arg	
65				70						75					80	
Val	Thr	Cys	Thr	Ala	Asp	Gly	Gly	His	Ser	Tyr	Pro	Gly	Ala	Leu	Pro	
			85						90					95		
Tyr	Asn	Ile	His	Ala	Arg	Leu	Glu	Arg	Gly	Val	Cys	Tyr	Asn	Gly	Trp	
			100					105					110			
Leu	Trp	Gly	Gly	Ala	Val	Asp	Asn									
		115				120										

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

233

- (A) LENGTH: 921 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(ix) FEATURE:

- (A) NAME/KEY: CDS
 (B) LOCATION: 1..921
 (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

ATG CTG CTC AGC CGT CAC AGG GAG CGC CTT GCC GCC AAC CTG GAG GAG	48
Met Leu Leu Ser Arg His Arg Glu Arg Leu Ala Ala Asn Leu Glu Glu	
1 5 10 15	
ACC GCC AAA GAC GCC GGA GAG AGG TGG GAA CTG AGT GCC CCG ACA TTC	96
Thr Ala Lys Asp Ala Gly Glu Arg Trp Glu Leu Ser Ala Pro Thr Phe	
20 25 30	
ACG CGA CAC TGT CCC AAA ACG GCA CGG ATG GCG CAC CCT TTT ATT GGC	144
Thr Arg His Cys Pro Lys Thr Ala Arg Met Ala His Pro Phe Ile Gly	
35 40 45	
GTG GTG CAC AGA ATA AAC TCA TAC AGT TCG GTC CTG GAA ACA TAC TGC	192
Val Val His Arg Ile Asn Ser Tyr Ser Ser Val Leu Glu Thr Tyr Cys	
50 55 60	
ACA CGG CAC CAT CCC GCC ACG CCC ACG TCA GCA AAT CCC GAC GTG GGA	240
Thr Arg His His Pro Ala Thr Pro Thr Ser Ala Asn Pro Asp Val Gly	
65 70 75 80	
ACC CCC AGA CCG TCC GAG GAC AAC GTC CCC GCA AAG CCG CGC CTA TTG	288
Thr Pro Arg Pro Ser Glu Asp Asn Val Pro Ala Lys Pro Arg Leu Leu	
85 90 95	
GAG TCC CTA TCA ACA TAC TTG CAG ATG CGG TGT GTG CGC GAG GAC GCG	336
Glu Ser Leu Ser Thr Tyr Leu Gln Met Arg Cys Val Arg Glu Asp Ala	
100 105 110	
CAC GTC TCC ACG GCC GAT CAA CTG GTC GAG TAC CAG GCG GGC AGA AAA	384
His Val Ser Thr Ala Asp Gln Leu Val Glu Tyr Gln Ala Gly Arg Lys	
115 120 125	
ACA CAC GAC TCC CTG CAC GCC TGC TCT GTC TAC CGC GAA CTT CAG GCT	432
Thr His Asp Ser Leu His Ala Cys Ser Val Tyr Arg Glu Leu Gln Ala	
130 135 140	
TTT CTG GTT AAC CTT TCG TCC TTT CTG AAC GGC TGT TAC GTT CCC GGG	480
Phe Leu Val Asn Leu Ser Ser Phe Leu Asn Gly Cys Tyr Val Pro Gly	
145 150 155 160	
GTG CAC TGG CTG GAG CCC TTC CAA CAG CAG CTA GTA ATG CAC ACT TTT	528
Val His Trp Leu Glu Pro Phe Gln Gln Leu Val Met His Thr Phe	
165 170 175	
TTC TTT TTG GTT TCA ATC AAG GCC CCA CAA AAG ACG CAC CAG TTG TTT	576
Phe Phe Leu Val Ser Ile Lys Ala Pro Gln Lys Thr His Gln Leu Phe	
180 185 190	

09607479.062900

234

GGA	TTG	TTT	AAG	CAG	TAC	TTC	GGT	TTA	TTT	GAA	ACT	CCA	AAC	AGT	GTT	624
Gly	Leu	Phe	Lys	Gln	Tyr	Phe	Gly	Leu	Phe	Glu	Thr	Pro	Asn	Ser	Val	
		195					200					205				
TTA	CAG	ACG	TTT	AAG	CAA	AAG	GCA	AGC	GTA	TTC	CTA	ATA	CCA	AGG	AGA	672
Leu	Gln	Thr	Phe	Lys	Gln	Lys	Ala	Ser	Val	Phe	Leu	Ile	Pro	Arg	Arg	
	210					215					220					
CAC	GGA	AAG	ACA	TGG	ATA	GTG	GTG	GCG	ATC	ATC	AGC	ATG	CTA	CTG	GCA	720
His	Gly	Lys	Thr	Trp	Ile	Val	Val	Ala	Ile	Ile	Ser	Met	Leu	Leu	Ala	
	225				230					235					240	
TCC	GTA	GAG	AAC	ATT	AAC	ATT	GGG	TAC	GTA	GCC	CAC	CAA	AAG	CAC	GTA	768
Ser	Val	Glu	Asn	Ile	Asn	Ile	Gly	Tyr	Val	Ala	His	Gln	Lys	His	Val	
				245				250						255		
GCC	AAC	TCC	GTG	TTC	GCG	GAA	ATC	ATA	AAG	ACG	CTT	TGT	CGG	TGG	TTC	816
Ala	Asn	Ser	Val	Phe	Ala	Glu	Ile	Ile	Lys	Thr	Leu	Cys	Arg	Trp	Phe	
			260				265						270			
CCC	CCC	AAA	AAT	TTA	AAC	ATC	AAG	AAG	GAG	AAC	GGA	ACC	ATA	ATC	TAC	864
Pro	Pro	Lys	Asn	Leu	Asn	Ile	Lys	Lys	Glu	Asn	Gly	Thr	Ile	Ile	Tyr	
		275					280					285				
ACG	CGA	CCC	GGA	GGA	CGG	TCC	AGC	TCG	CTG	ATG	TGC	GCA	ACA	TGC	TTC	912
Thr	Arg	Pro	Gly	Gly	Arg	Ser	Ser	Ser	Leu	Met	Cys	Ala	Thr	Cys	Phe	
	290					295					300					
AAT	AAG	AAC														921
Asn	Lys	Asn														
		305														

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 307 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

```

Met Leu Leu Ser Arg His Arg Glu Arg Leu Ala Ala Asn Leu Glu Glu
 1           5           10           15

Thr Ala Lys Asp Ala Gly Glu Arg Trp Glu Leu Ser Ala Pro Thr Phe
          20           25           30

Thr Arg His Cys Pro Lys Thr Ala Arg Met Ala His Pro Phe Ile Gly
          35           40           45

Val Val His Arg Ile Asn Ser Tyr Ser Ser Val Leu Glu Thr Tyr Cys
          50           55           60

Thr Arg His His Pro Ala Thr Pro Thr Ser Ala Asn Pro Asp Val Gly
          65           70           75           80

Thr Pro Arg Pro Ser Glu Asp Asn Val Pro Ala Lys Pro Arg Leu Leu
          85           90           95

Glu Ser Leu Ser Thr Tyr Leu Gln Met Arg Cys Val Arg Glu Asp Ala
          100           105           110

His Val Ser Thr Ala Asp Gln Leu Val Glu Tyr Gln Ala Gly Arg Lys
          115           120           125

```

235

Thr His Asp Ser Leu His Ala Cys Ser Val Tyr Arg Glu Leu Gln Ala
 130 135 140
 Phe Leu Val Asn Leu Ser Ser Phe Leu Asn Gly Cys Tyr Val Pro Gly
 145 150 155 160
 Val His Trp Leu Glu Pro Phe Gln Gln Gln Leu Val Met His Thr Phe
 165 170 175
 Phe Phe Leu Val Ser Ile Lys Ala Pro Gln Lys Thr His Gln Leu Phe
 180 185 190
 Gly Leu Phe Lys Gln Tyr Phe Gly Leu Phe Glu Thr Pro Asn Ser Val
 195 200 205
 Leu Gln Thr Phe Lys Gln Lys Ala Ser Val Phe Leu Ile Pro Arg Arg
 210 215 220
 His Gly Lys Thr Trp Ile Val Val Ala Ile Ile Ser Met Leu Leu Ala
 225 230 235 240
 Ser Val Glu Asn Ile Asn Ile Gly Tyr Val Ala His Gln Lys His Val
 245 250 255
 Ala Asn Ser Val Phe Ala Glu Ile Ile Lys Thr Leu Cys Arg Trp Phe
 260 265 270
 Pro Pro Lys Asn Leu Asn Ile Lys Lys Glu Asn Gly Thr Ile Ile Tyr
 275 280 285
 Thr Arg Pro Gly Gly Arg Ser Ser Ser Leu Met Cys Ala Thr Cys Phe
 290 295 300
 Asn Lys Asn
 305

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1365 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..1365
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

ATG GAT GCG CAT GCT ATC AAC GAA AGA TAC GTA GGT CCT CGC TGC CAC	48
Met Asp Ala His Ala Ile Asn Glu Arg Tyr Val Gly Pro Arg Cys His	
1 5 10 15	
CGT TTG GCC CAC GTG GTG CTG CCT AGG ACC TTT CTG CTG CAT CAC GCC	96
Arg Leu Ala His Val Val Leu Pro Arg Thr Phe Leu Leu His His Ala	
20 25 30	
ATA CCC CTG GAG CCC GAG ATC ATC TTT TCC ACC TAC ACC CGG TTC AGC	144

0960747-063000

[illegible][illegible]

237

TGT CGC ACG CAA GGC TAC ACG TTG AGA CAG CGC GTG CCT GTC GCC ATT 960
 Cys Arg Thr Gln Gly Tyr Thr Leu Arg Gln Arg Val Pro Val Ala Ile
 305 310 315 320

CCT CGC GAC GCG GAA ATC GCA GAC GCA GTT AAA TCG CAC TTT TTA GAG 1008
 Pro Arg Asp Ala Glu Ile Ala Asp Ala Val Lys Ser His Phe Leu Glu
 325 330 335

GCG TGC CTA GTG TTA CGG GGG CTG GCT TCG GAG GCT AGT GCC TGG ATA 1056
 Ala Cys Leu Val Leu Arg Gly Leu Ala Ser Glu Ala Ser Ala Trp Ile
 340 345 350

AGA GCT GCC ACG TCC CCG CCC CTT GGC CGC CAC GCC TGC TGG ATG GAC 1104
 Arg Ala Ala Thr Ser Pro Pro Leu Gly Arg His Ala Cys Trp Met Asp
 355 360 365

GTG TTA GGA TTA TGG GAA AGC CGC CCC CAC ACT CTA GGT TTG GAG TTA 1152
 Val Leu Gly Leu Trp Glu Ser Arg Pro His Thr Leu Gly Leu Glu Leu
 370 375 380

CGC GGC GTA AAC TGT GGC GGC ACG GAC GGT GAC TGG TTA GAG ATT TTA 1200
 Arg Gly Val Asn Cys Gly Gly Thr Asp Gly Asp Trp Leu Glu Ile Leu
 385 390 395 400

AAA CAG CCC GAT GTG CAA AAG ACA GTC AGC GGG AGT CTT GTG GCA TGC 1248
 Lys Gln Pro Asp Val Gln Lys Thr Val Ser Gly Ser Leu Val Ala Cys
 405 410 415

GTG ATC GTC ACA CCC GCA TTG GAA GCC TGG CTT GTG TTA CCT GGG GGT 1296
 Val Ile Val Thr Pro Ala Leu Glu Ala Trp Leu Val Leu Pro Gly Gly
 420 425 430

TTT GCT ATT AAA GCC CGC TAT AGG GCG TCG AAG GAG GAT CTG GTG TTC 1344
 Phe Ala Ile Lys Ala Arg Tyr Arg Ala Ser Lys Glu Asp Leu Val Phe
 435 440 445

ATT CGA GGC CGC TAT GGC TAG 1365
 Ile Arg Gly Arg Tyr Gly
 450

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 454 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

Met Asp Ala His Ala Ile Asn Glu Arg Tyr Val Gly Pro Arg Cys His
 1 5 10 15

Arg Leu Ala His Val Val Leu Pro Arg Thr Phe Leu Leu His His Ala
 20 25 30

Ile Pro Leu Glu Pro Glu Ile Ile Phe Ser Thr Tyr Thr Arg Phe Ser
 35 40 45

Arg Ser Pro Gly Ser Ser Arg Arg Leu Val Val Cys Gly Lys Arg Val
 50 55 60

Leu Pro Gly Glu Glu Asn Gln Leu Ala Ser Ser Pro Ser Gly Leu Ala
 65 70 75 80

006230-6240960

238

Leu Ser Leu Pro Leu Phe Ser His Asp Gly Asn Phe His Pro Phe Asp
 85 90 95
 Ile Ser Val Leu Arg Ile Ser Cys Pro Gly Ser Asn Leu Ser Leu Thr
 100 105 110
 Val Arg Phe Leu Tyr Leu Ser Leu Val Val Ala Met Gly Ala Gly Arg
 115 120 125
 Asn Asn Ala Arg Ser Pro Thr Val Asp Gly Val Ser Pro Pro Glu Gly
 130 135 140
 Ala Val Ala His Pro Leu Glu Glu Leu Gln Arg Leu Ala Arg Ala Thr
 145 150 155 160
 Pro Asp Pro Ala Leu Thr Arg Gly Pro Leu Gln Val Leu Thr Gly Leu
 165 170 175
 Leu Arg Ala Gly Ser Asp Gly Asp Arg Ala Thr His His Met Ala Leu
 180 185 190
 Glu Ala Pro Gly Thr Val Arg Gly Glu Ser Leu Asp Pro Pro Val Ser
 195 200 205
 Gln Lys Gly Pro Ala Arg Thr Arg His Arg Pro Pro Pro Val Arg Leu
 210 215 220
 Ser Phe Asn Pro Val Asn Ala Asp Val Pro Ala Thr Trp Arg Asp Ala
 225 230 235 240
 Thr Asn Val Tyr Ser Gly Ala Pro Tyr Tyr Val Cys Val Tyr Glu Arg
 245 250 255
 Gly Gly Arg Gln Glu Asp Asp Trp Leu Pro Ile Pro Leu Ser Phe Pro
 260 265 270
 Glu Glu Pro Val Pro Pro Pro Pro Gly Leu Val Phe Met Asp Asp Leu
 275 280 285
 Phe Ile Asn Thr Lys Gln Cys Asp Phe Val Asp Thr Leu Glu Ala Ala
 290 295 300
 Cys Arg Thr Gln Gly Tyr Thr Leu Arg Gln Arg Val Pro Val Ala Ile
 305 310 315 320
 Pro Arg Asp Ala Glu Ile Ala Asp Ala Val Lys Ser His Phe Leu Glu
 325 330 335
 Ala Cys Leu Val Leu Arg Gly Leu Ala Ser Glu Ala Ser Ala Trp Ile
 340 345 350
 Arg Ala Ala Thr Ser Pro Pro Leu Gly Arg His Ala Cys Trp Met Asp
 355 360 365
 Val Leu Gly Leu Trp Glu Ser Arg Pro His Thr Leu Gly Leu Glu Leu
 370 375 380
 Arg Gly Val Asn Cys Gly Gly Thr Asp Gly Asp Trp Leu Glu Ile Leu
 385 390 395 400
 Lys Gln Pro Asp Val Gln Lys Thr Val Ser Gly Ser Leu Val Ala Cys
 405 410 415
 Val Ile Val Thr Pro Ala Leu Glu Ala Trp Leu Val Leu Pro Gly Gly
 420 425 430
 Phe Ala Ile Lys Ala Arg Tyr Arg Ala Ser Lys Glu Asp Leu Val Phe

006699-6240360

445

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 984 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(ix) **FEATURE :**

(A) NAME/KEY: CDS

(B) LOCATION: 1..984

(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

ATG Met 1	TTT Phe	GCT Ala	TTG Leu	AGC Ser 5	TCG Ser	CTC Leu	GTG Val	TCC Ser	GAG Glu 10	GGT Gly	GAC Asp	CCG Pro	GAG Glu	GTG Val 15	ACC Thr	48
AGT Ser	AGG Arg	TAC Tyr	GTC Val 20	AAG Lys	GGC Gly	GTA Val	CAA Gln 25	CTT Leu	GCC Ala	CTG Leu	GAC Asp	CTT Leu	AGC Ser 30	GAG Glu	AAC Asn	96
ACA Thr	CCT Pro	GGA Gly 35	CAA Gln	TTT Phe	AAG Lys	TTG Leu	ATA Ile 40	GAA Glu	ACT Thr	CCC Pro	CTG Leu	AAC Asn 45	AGC Ser	TTC Phe	CTC Leu	144
TTG Leu	GTT Val 50	TCC Ser	AAC Asn	GTG Val	ATG Met	CCC Pro 55	GAG Glu	GTC Val	CAG Gln	CCA Pro	ATC Ile 60	TGC Cys	AGT Ser	GGC Gly	CGG Arg	192
CCG Pro 65	GCC Ala	TTG Leu	CGG Arg	CCA Pro	GAC Asp 70	TTT Phe	AGT Ser	AAT Asn	CTC Leu	CAC His 75	TTG Leu	CCT Pro	AGA Arg	CTG Leu	GAG Glu 80	240
AAG Lys	CTC Leu	CAG Gln	AGA Arg	GTC Val 85	CTC Leu	GGG Gly	CAG Gln	GGT Gly	TTC Phe 90	GGG Gly	GCG Ala	GCG Ala	GGT Gly	GAG Glu 95	GAA Glu	288
ATC Ile	GCA Ala	CTG Leu	GAC Asp 100	CCG Pro	TCT Ser	CAC His	GTA Val	GAA Glu 105	ACA Thr	CAC His	GAA Glu	AAG Lys	GGC Gly 110	CAG Gln	GTG Val	336
TTC Phe	TAC Tyr	AAC Asn 115	CAC His	TAT Tyr	GCT Ala	ACC Thr	GAG Glu 120	GAG Glu	TGG Trp	ACG Thr	TGG Trp	GCT Ala 125	TTG Leu	ACT Thr	CTG Leu	384
AAT Asn 130	AAG Lys	GAT Asp	GCG Ala	CTC Leu	CTT Leu	CGG Arg 135	GAG Glu	GCT Ala	GTA Val	GAT Asp	GGC Gly 140	CTG Leu	TGT Cys	GAC Asp	CCC Pro	432
GGA Gly 145	ACT Thr	TGG Trp	AAG Lys	GGT Gly	CTT Leu 150	CTT Leu	CCT Pro	GAC Asp	GAC Asp	CCC Pro 155	CTT Leu	CCG Pro	TTG Leu	CTA Leu	TGG Trp 160	480

240

CTG	CTG	TTC	AAC	GGA	CCC	GCC	TCT	TTT	TGT	CGG	GCC	GAC	TGT	TGC	CTG	528
Leu	Leu	Phe	Asn	Gly	Pro	Ala	Ser	Phe	Cys	Arg	Ala	Asp	Cys	Cys	Leu	
			165						170					175		
TAC	AAG	CAG	CAC	TGC	GGT	TAC	CCG	GGC	CCG	GTG	CTA	CTT	CCA	GGT	CAC	576
Tyr	Lys	Gln	His	Cys	Gly	Tyr	Pro	Gly	Pro	Val	Leu	Leu	Pro	Gly	His	
			180					185					190			
ATG	TAC	GCT	CCC	AAA	CGG	GAT	CTT	TTG	TCG	TTC	GTT	AAT	CAT	GCC	CTG	624
Met	Tyr	Ala	Pro	Lys	Arg	Asp	Leu	Leu	Ser	Phe	Val	Asn	His	Ala	Leu	
		195					200					205				
AAG	TAC	ACC	AAG	TTT	CTA	TAC	GGA	GAT	TTT	TCC	GGG	ACA	TGG	GCG	GCG	672
Lys	Tyr	Thr	Lys	Phe	Leu	Tyr	Gly	Asp	Phe	Ser	Gly	Thr	Trp	Ala	Ala	
	210					215					220					
GCT	TGC	CGC	CCG	CCA	TTC	GCT	ACT	TCT	CGG	ATA	CAA	AGG	GTA	GTG	AGT	720
Ala	Cys	Arg	Pro	Pro	Phe	Ala	Thr	Ser	Arg	Ile	Gln	Arg	Val	Val	Ser	
225					230					235					240	
CAG	ATG	AAA	ATC	ATA	GAT	GCT	TCC	GAC	ACT	TAC	ATT	TCC	CAC	ACC	TGC	768
Gln	Met	Lys	Ile	Ile	Asp	Ala	Ser	Asp	Thr	Tyr	Ile	Ser	His	Thr	Cys	
			245					250						255		
CTC	TTG	TGT	CAC	ATA	TAT	CAG	CAA	AAT	AGC	ATA	ATT	GCG	GGT	CAG	GGG	816
Leu	Leu	Cys	His	Ile	Tyr	Gln	Gln	Asn	Ser	Ile	Ile	Ala	Gly	Gln	Gly	
			260					265					270			
ACC	CAC	GTG	GGT	GGA	ATC	CTA	CTG	TTG	AGT	GGA	AAA	GGG	ACC	CAG	TAT	864
Thr	His	Val	Gly	Gly	Ile	Leu	Leu	Leu	Ser	Gly	Lys	Gly	Thr	Gln	Tyr	
		275					280					285				
ATA	ACA	GGC	AAT	GTT	CAG	ACC	CAA	AGG	TGT	CCA	ACT	ACG	GGC	GAC	TAT	912
Ile	Thr	Gly	Asn	Val	Gln	Gln	Arg	Cys	Pro	Thr	Thr	Gly	Asp	Tyr		
	290					295				300						
CTA	ATC	ATC	CCA	TCG	TAT	GAC	ATA	CCG	GCG	ATC	ATC	ACC	ATG	ATC	AAG	960
Leu	Ile	Ile	Pro	Ser	Tyr	Asp	Ile	Pro	Ala	Ile	Ile	Thr	Met	Ile	Lys	
305					310					315					320	
GAG	AAT	GGA	CTC	AAC	CAA	CTC	TAA									984
Glu	Asn	Gly	Leu	Asn	Gln	Leu										
				325												

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 327 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

Met	Phe	Ala	Leu	Ser	Ser	Leu	Val	Ser	Glu	Gly	Asp	Pro	Glu	Val	Thr	
1				5					10					15		
Ser	Arg	Tyr	Val	Lys	Gly	Val	Gln	Leu	Ala	Leu	Asp	Leu	Ser	Glu	Asn	
			20					25					30			
Thr	Pro	Gly	Gln	Phe	Lys	Leu	Ile	Glu	Thr	Pro	Leu	Asn	Ser	Phe	Leu	
		35					40					45				
Leu	Val	Ser	Asn	Val	Met	Pro	Glu	Val	Gln	Pro	Ile	Cys	Ser	Gly	Arg	
	50					55					60					

241

Pro Ala Leu Arg Pro Asp Phe Ser Asn Leu His Leu Pro Arg Leu Glu
 65 70 75 80
 Lys Leu Gln Arg Val Leu Gly Gln Gly Phe Gly Ala Ala Gly Glu Glu
 85 90 95
 Ile Ala Leu Asp Pro Ser His Val Glu Thr His Glu Lys Gly Gln Val
 100 105 110
 Phe Tyr Asn His Tyr Ala Thr Glu Glu Trp Thr Trp Ala Leu Thr Leu
 115 120 125
 Asn Lys Asp Ala Leu Leu Arg Glu Ala Val Asp Gly Leu Cys Asp Pro
 130 135 140
 Gly Thr Trp Lys Gly Leu Leu Pro Asp Asp Pro Leu Pro Leu Leu Trp
 145 150 155 160
 Leu Leu Phe Asn Gly Pro Ala Ser Phe Cys Arg Ala Asp Cys Cys Leu
 165 170 175
 Tyr Lys Gln His Cys Gly Tyr Pro Gly Pro Val Leu Leu Pro Gly His
 180 185 190
 Met Tyr Ala Pro Lys Arg Asp Leu Leu Ser Phe Val Asn His Ala Leu
 195 200 205
 Lys Tyr Thr Lys Phe Leu Tyr Gly Asp Phe Ser Gly Thr Trp Ala Ala
 210 215 220
 Ala Cys Arg Pro Pro Phe Ala Thr Ser Arg Ile Gln Arg Val Val Ser
 225 230 235 240
 Gln Met Lys Ile Ile Asp Ala Ser Asp Thr Tyr Ile Ser His Thr Cys
 245 250 255
 Leu Leu Cys His Ile Tyr Gln Gln Asn Ser Ile Ile Ala Gly Gln Gly
 260 265 270
 Thr His Val Gly Gly Ile Leu Leu Leu Ser Gly Lys Gly Thr Gln Tyr
 275 280 285
 Ile Thr Gly Asn Val Gln Thr Gln Arg Cys Pro Thr Thr Gly Asp Tyr
 290 295 300
 Leu Ile Ile Pro Ser Tyr Asp Ile Pro Ala Ile Ile Thr Met Ile Lys
 305 310 315 320
 Glu Asn Gly Leu Asn Gln Leu
 325

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 330 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

242

GGATCCCTCT GACAACCTTC AGATAAAAAA CGTATATGCC CCCTTTTTTTC AGTGGGACAG 60
CAACACCCAG CTAGCAGTGC TACCCCGATT TTTTAGCCGA AAGGATTCCA CCATTGTGCT 120
CGAATCCAAC GGATTTGACC CCGTGTTOCC CATGGTCGTG CCGCAGCAAC TGGGGCACGC 180
TATTCTGCAG CAGCTGTTGG TGTACCACAT CTACTCCAAA ATATCGGCCG GGGCCCCGGA 240
TGATGTAAAT ATGGCGGAAC TTGATCTATA TACCACCAAT GTGTCATTTA TGGGGCGCAC 300
ATATCGTCTG GACGTAGACA ACACGGATCC 330

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 627 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

GGATCCGCTG GCAGGTGGGC GCGCACCTCG TCGGGTAGCT TGGAGACAAA CAGCTCCAGG 60
CCAGTCCGCG CCGTAGCGCC TGCAGGTGCC TCACCACCGG GGCCGGGTCA TGCATCTGT 120
TTAGTCCGGA GAAGATAGGG CCCTTGGGAA GCCGCTGARC CAGCTCCAGG GTCTCCAAGA 180
TGCGCACCGG TTGTCGGAGC TGTCGCGATA GAGGTTAGGG TAGGTGTCCG GTCCGTCCGT 240
GGGCTCAAAC CTGCCCAGAC ACACCACTGT CTGCTGGGGG ATCATCCTTC TCAGGGAGAT 300
GCATTCTTTG GAAGTAGTGG TAGAGATGGA GCAGACTGCC AGGGCGTTGC AGGAGTGGTG 360
GCGATGGTGC GCACCGTTTT TAAGAAACCC CCCAGGGTGG GCACTCCCGC TCCCTGCAGC 420
ATCTCGGCCT GCTGTACGTC CTTGGCGAAT ATGCGACGAA ATGGGCTGTG CGCACGGGGT 480
CCCAGGGCCG GTCCGGTGGC ATACAGGCCG GTGAGGGCCC CCTGGGTCTG TCCGCTTGA 540
AACAGGGTGC TGTGAAACAA CAGGTTGCAA GGCCGCGAAT ACCCTCTGTC ACGTGCTGT 600
GGACGTGGGT GTATGCTCCG TGGATCC 627

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 233 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

243

AGCCGAAAGG ATTCCACCAT TGTGCTCGAA TCCAACGGAT TTGACCCCGT GTTCCCCATG 60
 GTCGTGCCGC AGCAACTGGG GCACGCTATT CTGCAGCAGC TGTTGGTGTA CCACATCTAC 120
 TCCAAAATAT CGGCCGGGGC CCCGGATGAT GTAAATATGG CGGAACTTGA TCTATATACC 180
 ACCAATGTGT CATTATATGGG GCGCACATAT CGTCTGGACG TAGACAACAC GGA 233

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 328 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

GAAATTACCC ACGAGATCGC TTCCCTGCAC ACCGCACTTG GCTACTCATC AGTEATCGCC 60
 CCGGCCACG TGGCCGCCAT AACTACAGAC ATGGGAGTAC ATTGTCAGGA CCTCTTTATG 120
 ATTTTCCCAG GGGACGCGTA TCAGGACCGC CAGCTGCAATG ACTATATCAA AATGAAAGCG 180
 GGCGTGCAA CCGGCTCACC GGGAAACAGA ATGGATCAGC TGGGATACAC TGCTGGGGTT 240
 CCTCGCTGCG AGAACCTGCC CGGTTTGAGT CATGGTCAGC TGGCAACCTG CGAGATAATT 300
 CCCACGCCGG TCACATCTGA CGTTGCCT 328

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 132 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

AACACGTCAT GTGCAGGAGT GACATTGTGC CGCGGAGAAA CTCAGATCGC ATCCCGTAA 60
 CACACTGAGT GGGAAAATCT GCTGGCTATG TTTTCTGTGA TTATCTATGC CTTAGATCAC 120
 AACTGTCACC CG 132

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 40 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single

0960740960

244

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

AGCCGAAAGG ATTCCACCAT TCCGTGTTGT CTACGTCCAG

40

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 38 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

GAAATTACCC ACGAGATCGC AGGCAACGTC AGATGTGA

38

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 46 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

AACACGTCAT GTGCAGGAGT GACCGGGTGA CAGTTGTGAT CTAAGG

46

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

